

1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_nam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1192	100.8	1192	6	AX431088	AX431088 Sequence
2	1082.4	90.0	1128	9	BC013587	BC013587 Homo sapi
3	1012	84.9	1049	8	BD191093	BD191093 166 human
4	808.8	67.9	165110	9	AL450311	AL450311 Human DNA
c	801.6	67.2	173341	2	AC021954	AC021954 Homo sapi
6	473	39.7	483	6	AR413227	AR413227 Sequence
7	473	39.7	483	6	BD108780	BD108780 EST and e
8	452.6	38.0	550	6	AX131082	AX131082 Sequence
9	432.8	36.3	643	6	AX431087	AX431087 Sequence
10	391.8	32.9	519	6	AR415427	AR415427 Sequence
11	391.8	32.9	519	6	BD110980	BD110980 EST and e
c	375.4	31.5	480	11	G22561	G22561 human STS W
13	307.4	25.8	1538	10	BC024943	BC024943 Mus muscu
c	306	25.7	275924	2	AC116233	AC116233 Rattus no
14	294	24.7	240657	2	AC116233	AC116233 Rattus no
c	294	21.0	258815	2	AC127817	AC127817 Rattus no
15	250.8	21.0	210500	2	AC127417	AC127417 Mus muscu
c	250.8	21.0	210500	2	AC127417	AC127417 Mus muscu
17	236.4	19.8	138070	2	AR425705	AR425705 Sequence
c	236.4	19.8	138070	2	AR425705	AR425705 Sequence
19	233.8	19.5	364	6	BD121258	BD121258 EST and e
20	233.8	19.5	364	6	AC099492	AC099492 Rattus no
c	233.8	19.5	364	6	AC099492	AC099492 Rattus no
21	88.2	7.4	238937	2	AL111858	AL111858 Mus muscu
22	75.4	6.3	177226	10	AL590988	AL590988 Mouse DNA
23	75.4	6.3	183975	6	AX644366	AX644366 Sequence
24	75	6.3	986	6	AX535036	AX535036 Sequence
25	69.6	5.8	577	6	AX644276	AX644276 Sequence
26	65.2	5.5	737	6	AX431080	AX431080 Sequence
27	64	5.4	666	6	AX431141	AX431141 Sequence
28	63.6	5.3	626	6	AX644313	AX644313 Sequence
29	60.4	5.2	879	6	AX535007	AX535007 Sequence
30	60.4	5.1	1545	6	AX497809	AX497809 Sequence
31	58.8	5.0	1623	6	AX431074	AX431074 Sequence
32	58.4	4.9	510	6	AX431097	AX431097 Sequence
33	57.8	4.8	2910	6	PFMAL6P1.10	PFMAL6P1.10 Continuation (11 o
c	57.6	4.8	110000	2	AC117075	AC117075 Dictyoste
34	57.4	4.8	254733	3	AC116978	AC116978 Dictyoste
c	57.4	4.8	27785	2	AC117076	AC117076 Dictyoste
36	57.2	4.8	183648	3	AE014836	AE014836 Plasmodu
c	57.2	4.8	250743	3	AE014836	AE014836 Plasmodu
38	57.2	4.8	250743	3	AE014836	AE014836 Plasmodu
c	57.2	4.8	250743	3	AE014836	AE014836 Plasmodu
39	57.2	4.8	250743	3	AE014836	AE014836 Plasmodu
c	57.2	4.8	250743	3	AE014836	AE014836 Plasmodu
40	56.8	4.8	200846	2	AE0146514	AE0146514 Pan trogl
c	56.8	4.8	200846	2	AE0146514	AE0146514 Pan trogl
41	56.6	4.7	250029	3	RRH45STU	RRH45STU R.rattus mR
42	56.4	4.7	1066	10	AC116982	AC116982 Dictyoste
c	56.4	4.7	256879	3	AC116982	AC116982 Dictyoste
43	56.4	4.7	256879	3	AC116982	AC116982 Dictyoste
c	56.4	4.7	256879	3	AC116982	AC116982 Dictyoste
44	55.6	4.7	182871	3	AC117176	AC117176 Dictyoste
c	55.6	4.7	182871	3	AC117176	AC117176 Dictyoste

ALIGNMENTS

RESULT 1					
AX431088					
LOCUS	AX431088	1192 bp.	DNA	linear	PAT 28-JUN-2002
DEFINITION	Sequence 31 from Patent WO0240535.				
ACCESSION	AX431088				
VERSION	AX431088.1	GI:21656065			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1				
AUTHORS	Reippon, H., Liu, C., Salceda, S., Sun, Y., Cafferty, R. and				
	Macina, R.A.				
TITLE	Compositions and methods relating to ovary specific genes and				

Proteins
 Patent: WO 0240535-A 31 23-MAY-2002;
 DIADEXUS INC (US)
 Location/Qualifiers
 1..1192
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1192; DB 6; Length 1192;
 Best Local Similarity 100.0%; Pred. No. 1.7e-288;
 Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAACCGCTGAGAGCGCGGAGTGAAGCGCGCCGAGCGCGGAGTGCAGCTGACACA 60
DB 1 AAACCGCTGAGAGCGCGGAGTGAAGCGCGCCGAGCGCGGAGTGCAGCTGACACA 60
QY 61 CAGCATCTACTAGAGGTGGGTCACTCTGTAAACATCACTGACTGCAAGCTTCCCTGAT 120
DB 61 CAGCATCTACTAGAGGTGGGTCACTCTGTAAACATCACTGACTGCAAGCTTCCCTGAT 120
QY 121 TTCTGGTGAAGCCCATCAAGGACCCAGCGCTGAGAGATGCTGCGGATCTTGCCCAA 180
DB 121 TTCTGGTGAAGCCCATCAAGGACCCAGCGCTGAGAGATGCTGCGGATCTTGCCCAA 180
QY 181 TGGGGAATTCGTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 TGGGGAATTCGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 TAGCATTCCTCGACAGAGCTTCTTCATATAGGGGCCATGCTCTCCCGAGGGGTCTCGG 300
DB 241 TAGCATTCCTCGACAGAGCTTCTTCATATAGGGGCCATGCTCTCCCGAGGGGTCTCGG 300
QY 301 CCCCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 360
DB 301 CCCCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 360
QY 361 CAACCGGAGCTGTGAGAGATGGGCTTCCGAGTGGAGATCTCGGCAACATGCTGTGGA 420
DB 361 CAACCGGAGCTGTGAGAGATGGGCTTCCGAGTGGAGATCTCGGCAACATGCTGTGGA 420
QY 421 GCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
DB 421 GCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 481 CCGTGTGGGCTTGTCTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
DB 481 CCGTGTGGGCTTGTCTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 541 TAGGGGTGGGTTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 TAGGGGTGGGTTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 AGGGGATCTGGGTGGCTTGAAGATGATCAAGAGGGGACCAAGGTGTGTGTGTTC 660
DB 601 AGGGGATCTGGGTGGCTTGAAGATGATCAAGAGGGGACCAAGGTGTGTGTGTTC 660
QY 661 CTTTGTGTAAAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 CTTTGTGTAAAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 GGTCCGTGTGTCCGAGATCTTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 721 GGTCCGTGTGTCCGAGATCTTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 781 TCAAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 840
DB 781 TCAAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 840
QY 841 CAGCTGTAGTCCAAATCTTCGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 CAGCTGTAGTCCAAATCTTCGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

```

```

QY 901 GGGAGCAGAGACTTGGGCTTAGGGCAGGTGAAAAATTCAGACTTTTATGACATGTTT 960
DB 901 GGGAGCAGAGACTTGGGCTTAGGGCAGGTGAAAAATTCAGACTTTTATGACATGTTT 960
QY 961 TTGTTTAAATGATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1020
DB 961 TTGTTTAAATGATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1020
QY 1021 TCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1080
DB 1021 TCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1080
QY 1081 AAAAAGCTGTGGGGTTTACCCGGGGCCAAAGGGGGGCCCGGGGGGGGAATGTGCTTCC 1140
DB 1081 AAAAAGCTGTGGGGTTTACCCGGGGCCAAAGGGGGGCCCGGGGGGGGAATGTGCTTCC 1140
QY 1141 CGCCCAATTCCTCCCAATTTTGTACAAATGAGAGAGACACGGAATCC 1192
DB 1141 CGCCCAATTCCTCCCAATTTTGTACAAATGAGAGAGACACGGAATCC 1192

```

RESULT 2
 BC013587
 LOCUS
 DEFINITION
 Homo sapiens chromosome 10 open reading frame 35, mRNA (cDNA clone MGC:9536 IMAGE:3896656), complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tomshy, L.S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richardson, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1129)
 JOURNAL
 PUBMED
 STRAUSBERG, R.
 DIRECT SUBMISSION
 Submitted (04-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Array at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAC Plate: 14 Row: 1 Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21687152.

FEATURES

source

Location/Qualifiers

1. 1129
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:9596 IMAGS:3896656"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_id="NH_MGC_70"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1129
/gene="C10orf35"
/db_xref="LocusID:219738"
1607. 525
/codon_start=1
/product="chromosome 10 open reading frame 35"
/protein_id="AAH13587.1"
/db_xref="GI:15488920"
/db_xref="LocusID:219738"
/translation="VRIILANGBIIVDDPRVTTTTPRGSTIPRSFNRHGADPG
GPRGQQAGARLGAASFPNDLNRDLVNMSPFQWMLGNHVAEPTVSTILLFLMLMG
VRGLLVGLVYLVSHLSOR"

CDS

gene

ORIGIN

Query Match 90.8%; Score 1082.4; DB 9; Length 1129;
Best Local Similarity 99.9%; Pred. No. 6e-261; Indels 0; Gaps 0;
Matches 1083; Conservative 0; Mismatches 1;

QY 2 AACCGCTGAGCCGCCGAGAGTGAAGCGCCGCGAGGCCGAGAGTGCCTGTCAGACAC 61
Db 1 AACCGCTGAGCCGCCGAGAGTGAAGCGCCGCGAGGCCGAGAGTGCCTGTCAGACAC 60
QY 62 AGCATCTACTCAGAGTGGTCACTCTGTGAACATCATCTGACAGCTCCCTCAATT 121
Db 61 AGCATCTACTCAGAGTGGTCACTCTGTGAACATCATCTGACAGCTCCCTCAATT 120
QY 122 TCTGTGACAGCCCATCAAGGACCAAGCCCTGAGAGATGGTGGCATCTTGGCCAAAT 181
Db 121 TCTGTGACAGCCCATCAAGGACCAAGCCCTGAGAGATGGTGGCATCTTGGCCAAAT 180
QY 182 GGGGAAATCTGTGACAGACGACCCCGAGTGAAGCACTAACCCAGCCACCAAGAGT 241
Db 181 GGGGAAATCTGTGACAGACGACCCCGAGTGAAGCACTAACCCAGCCACCAAGAGT 240
QY 242 AGCATCTCTGACAGACCTCTTCAATAGGGGCCATGTGCTCCCGCAGGGGCTCTGGC 301
Db 241 AGCATCTCTGACAGACCTCTTCAATAGGGGCCATGTGCTCCCGCAGGGGCTCTGGC 300
QY 302 CCCCGCAGAGCAGGAGGTGCAGGCTGGTGTCTGCTCAATCCCTTCAATGACCTC 361
Db 301 CCCCGCAGAGCAGGAGGTGCAGGCTGGTGTCTGCTCAATCCCTTCAATGACCTC 360
QY 362 AACCGGAGAGTGTGAACATGGGCTTTCGGCAGTGGATCTGGGCAACATGCTGTGAG 421
Db 361 AACCGGAGAGTGTGAACATGGGCTTTCGGCAGTGGATCTGGGCAACATGCTGTGAG 420
QY 422 CCGGTGACCTCATCTGCTCTCTTCTGCTCATGATGCTGTGGTGTGGTGGCTCTGC 481
Db 421 CCGGTGACCTCATCTGCTCTCTTCTGCTCATGATGCTGTGGTGTGGTGGCTCTGC 480
QY 482 CTGGTGGCTTGTCTAAGTGTGTCCACCTGATGACGCGTGAAGCTCTGAGGGCTGAT 541
Db 481 CTGGTGGCTTGTCTAAGTGTGTCCACCTGATGACGCGTGAAGCTCTGAGGGCTGAT 540

QY 542 AGGGGTGGTGTGTGTGAAGGAGCTTGTGCGGGCTTGTGTGAGAGCAGGCATATTTGGA 601
Db 541 AGGGGTGGTGTGTGTGAAGGAGCTTGTGCGGGCTTGTGTGAGAGCAGGCATATTTGGA 600
QY 602 GGGATCTGTGTGTCCTTGAAGTATGATCAGAGAGGAGCAGAGGTGTGTGTCC 661
Db 601 GGGATCTGTGTGTCCTTGAAGTATGATCAGAGAGGAGCAGAGGTGTGTGTCC 660
QY 662 CTTTGTGTTAAGCTGAGGAGAGGAGAGCTTATGTCAGATTTCCAAAGTGTGGTGG 721
Db 661 CTTTGTGTTAAGCTGAGGAGAGGAGAGCTTATGTCAGATTTCCAAAGTGTGGTGG 720
QY 722 GTCCTGTGTTCCGAGATCTTTAGAGTGAATGGGGCTGTCATTAAGTGGCAAAAT 781
Db 721 GTCCTGTGTTCCGAGATCTTTAGAGTGAATGGGGCTGTCATTAAGTGGCAAAAT 780
QY 782 CAGAGCAAGAAAGCAGATGCTTCCCAATCTCTCAATCTTTTAGCCAGAGATCTC 841
Db 781 CAGAGCAAGAAAGCAGATGCTTCCCAATCTCTCAATCTTTTAGCCAGAGATCTC 840
QY 842 AGCTGATGCCCAACATGTTCCGATGCTGTGAACATGCCAGAGTCTCTCTGCTAG 901
Db 841 AGCTGATGCCCAACATGTTCCGATGCTGTGAACATGCCAGAGTCTCTCTGCTAG 900
QY 902 GAGAGAGACTTGGGCTTAGGCGAGAGTGAAGAAATTCAGACTTTTGAAGACTGTTT 961
Db 901 GAGAGAGACTTGGGCTTAGGCGAGAGTGAAGAAATTCAGACTTTTGAAGACTGTTT 960
QY 962 TGTGTTAATGTAATTTTATGAGCTTATTTAGTGTGAGCAAGTGTAGTGGCAAT 1021
Db 961 TGTGTTAATGTAATTTTATGAGCTTATTTAGTGTGAGCAAGTGTAGTGGCAAT 1020
QY 1022 CTAATTAATGTAACCTTTCAATAATAGATTAGTAATTAATTAATTAATTAATTA 1081
Db 1021 CTAATTAATGTAACCTTTCAATAATAGATTAGTAATTAATTAATTAATTAATTA 1080
QY 1082 AAAA 1085
Db 1081 AAAA 1084
RESULT 3
BD191093
LOCUS
DEFINITION
186 human secreted proteins.
ACCESSION
BD191093.1 GI:33000832
VERSION
JP 2002510192-A/57.
KEYWORDS
unidentified
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1049)
Ruben,S.M., Rosen,C.A., Fischer,C.L., Soppet,D.R., Carter,K.C.,
Bednarik,D.P., Endress,G.A., Yu,G.L., N.J., Feng,P., Young,P.E.,
Greene,J.M., Fertle,A.M., Duan,R., Hu,U.S., Florence,K.A.,
Olson,H.S., Eder,R., Brewer,L.A., Moore,P.A., Shi,Y.,
Lafleur,D.W., Li,Y., Zeng,Z. and Kyaw,H.
186 human secreted proteins
Patent: JP 2002510192-A 57 02-APR-2002;
HUMAN GENOME SCIENCES INC
PN JP 2002510192-A/57
COMMENT
PD 02-APR-2002
PF 06-MAR-1998 JP 1998538833
PR 07-MAR-1997 US 60/040162, 07-MAR-1997 US 60/040333 PR
07-MAR-1997 US 60/038621, 07-MAR-1997 US 60/040161 PR
07-MAR-1997 US 60/040626, 07-MAR-1997 US 60/040334 PR
07-MAR-1997 US 60/040336, 07-MAR-1997 US 60/040163 PR
11-APR-1997 US 60/043580, 11-APR-1997 US 60/043568 PI
M RUBEN, CRAIG A ROSEN, KARRIE L FISCHER, DANIEL R SOPPET, PI
KENNETH C CARTER, DANIEL P BEDNARIK, GREGORY
A ENDRESS, GUO LIANG
PI YU, JIAN NI,
PI PING FENG, PAUL E YOUNG, JOHN M GREENE, ANN

M FERRIE, ROXANNE DUAN,
 PI JING SHAN HU, KIMBERLY A FLORENCE, HENRIK
 S OLSEN, REINHARD EBERER,
 PI LAURIE A BREWER, PAUL A MOORE, YANGU SHI, DAVID W LAFLEUR PI
 , YI LI, ZHI ZHEN ZENG,
 PI HLA KYAM
 PC C12N15/12, C12N5/10, C12N1/21, C07K14/47, C07K16/18, C12Q1/68, PC
 G01N33/50,
 PC G01N33/53, G01N33/68, A61K38/17
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers.

FEATURES
 source 1..1049
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

Query Match 84.9%; Score 1012; DB 6; Length 1049;
 Best Local Similarity 99.3%; Pred. No. 3e-243;
 Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

46 TCGGCGCTGCGACACAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACTG 105
 1 TCGGCGCTGCGACACAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACTG 60
 106 CAAGCTCTCCCTCAATTTCTGTGTGAGGCCCATCAGAGGACCCACAGCGCTGGAGAGTGT 165
 61 CAAGCTCTCCCTCAATTTCTGTGTGAGGCCCATCAGAGGACCCACAGCGCTGGAGAGTGT 120
 166 GCGGATCTTGGCCCAATGGGAAATGTGTGACAGACAGACGCCCGAGTGTGAGACCTAC 225
 121 GCGGATCTTGGCCCAATGGGAAATGTGTGACAGACAGACGCCCGAGTGTGAGACCTAC 180
 226 CGAGGACACAGAGGTGACATTCCTCCAGACAGAGCTTCTCAATGAGGCGCATGCTGCTC 285
 181 CGAGGACACAGAGGTGACATTCCTCCAGACAGAGCTTCTCAATGAGGCGCATGCTGCTC 240
 286 CCGAGGCGGTCTGTGCGCCCGCCGACAGACAGAGGTGCGAGGCTGGGTGCTGCTCAGTC 345
 241 CCGAGGCGGTCTGTGCGCCCGCCGACAGACAGAGGTGCGAGGCTGGGTGCTGCTCAGTC 300
 346 CCGCTTCAATATACCTCAACCGGCGAGCTGTGAGACATGGGCTTCCGAGTGGGACATCTCG 405
 301 CCGCTTCAATATACCTCAACCGGCGAGCTGTGAGACATGGGCTTCCGAGTGGGACATCTCG 360
 406 CAACCATCTGTGAGAGCGGTGACCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 465
 361 CAACCATCTGTGAGAGCGGTGACCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
 466 TGTTCGTGGCT 525
 421 TGTTCGTGGCT 480
 526 ACGCTGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGG 585
 481 ACGCTGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGG 540
 586 AGCAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGG 645
 541 AGCAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGGTGAGAGGCGG 600
 646 CAGGTGTGTGTCT 703
 601 CAGGTGTGTGTCT 660
 704 TTTCCAAAGT 763
 661 TTTCCAAAGT 720
 764 CATTAAGT 823

DB 721 CATTAAGT 780
 QY 824 TTAATCCGAGAGAGATCTCAGCTGATGACCAATGTTCCGATGCTGTGAGAGATGCC 883
 DB 781 TTAATCCGAGAGAGATCTCAGCTGATGACCAATGTTCCGATGCTGTGAGAGATGCC 840
 QY 884 GAGCT 943
 DB 841 GAGCT 900
 QY 944 CTTTCTTGAACCTGTCT 1003
 DB 901 CTTTCTTGAACCTGTCT 960
 QY 1004 ACAAGT 1063
 DB 961 ACAAGT 1020
 QY 1064 AAAAAAAAAAAAAA 1077
 DB 1021 AAAAAAAAAAAAAA 1034

RESULT 4
 AL450311
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 165110)
 Howden, P.
 Direct Submission
 Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jul 8, 2001 this sequence version replaced gi:14575291.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Emi, EMBL; Sw,
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr10
 RP11-343J3 is from the library RPc1-11.2 constructed by the group
 of Pletzer de Jong. For further details see
 http://www.chori.org/bacpac/home.htm

VECTOR: pBACe3.6
 This sequence is the entire insert of clone RP11-343J3 The true
 left end of clone RP11-242620 is at 139955 in this sequence. The
 true right end of clone RP11-40466 is at 6588 in this sequence.
 Location/Qualifiers
 1..165110

FEATURES
 source

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-343J3"
/clone_lib="RP11-11.2"
7. .147
repeat_region /note="THE1C repeat: matches 2. .142 of consensus"
151. .293
repeat_region /note="MIR repeat: matches 2. .148 of consensus"
1136. .1187
/note="26 copies 2 mer gt 98% conserved"
2960. .3272
repeat_region /note="AluSg repeat: matches 1. .310 of consensus"
4144. .4270
/note="AluSx repeat: matches 1. .134 of consensus"
4319. .4494
/note="AluSx repeat: matches 118. .293 of consensus"
5662. .6438
repeat_region /note="L1MC repeat: matches 272. .1095 of consensus"
6916. .8080
/note="L1MC repeat: matches 1168. .2367 of consensus"
9934. .10087
repeat_region /note="MIR repeat: matches 25. .184 of consensus"
10353. .10481
/note="MIR repeat: matches 65. .194 of consensus"
11077. .11372
repeat_region /note="AluSx repeat: matches 1. .294 of consensus"
11600. .11639
/note="20 copies 2 mer tc 95% conserved"
12479. .12608
repeat_region /note="65 copies 2 mer at 90% conserved"
13433. .13876
/note="MTR1C repeat: matches 1. .466 of consensus"
14372. .14576
/note="AluSg repeat: matches 1. .301 of consensus"
14577. .14633
repeat_region /note="AluSg repeat: matches 243. .299 of consensus"
14843. .14943
repeat_region /note="MER86 repeat: matches 9. .109 of consensus"
15404. .15613
/note="MIR repeat: matches 3. .213 of consensus"
16626. .16705
repeat_region /note="MIR repeat: matches 60. .140 of consensus"
16928. .16993
/note="33 copies 2 mer gg 66% conserved"
17476. .17569
repeat_region /note="MER81 repeat: matches 2. .114 of consensus"
17719. .18069
/note="L1MC4 repeat: matches 7617. .7977 of consensus"
18117. .18345
repeat_region /note="MIR repeat: matches 7. .262 of consensus"
19169. .19400
repeat_region /note="MIR repeat: matches 5. .239 of consensus"
19436. .19627
/note="L1MC5 repeat: matches 7728. .7917 of consensus"
19642. .19716
repeat_region /note="MIR repeat: matches 48. .131 of consensus"
19969. .20010
/note="21 copies 2 mer tg 100% conserved"
20623. .20704
repeat_region /note="41 copies 2 mer gt 85% conserved"
20738. .20773
/note="U2 repeat: matches 1. .36 of consensus"
20802. .20863
repeat_region /note="MTR1 repeat: matches 1. .62 of consensus"
20889. .21263
repeat_region /note="MTR1P repeat: matches 188. .541 of consensus"
21463. .21618
/note="MIR repeat: matches 46. .192 of consensus"
22019. .22326
repeat_region /note="AluB repeat: matches 1. .306 of consensus"
22381. .22564

repeat_region /note="U2 repeat: matches 2453. .2629 of consensus"
22896. .23174
repeat_region /note="MTR1D repeat: matches 117. .413 of consensus"
23215. .23346
/note="MIR repeat: matches 48. .188 of consensus"
23388. .23532
repeat_region /note="U2 repeat: matches 2097. .2230 of consensus"
23533. .23837
repeat_region /note="AluB repeat: matches 1. .303 of consensus"
23838. .24137
repeat_region /note="U2 repeat: matches 1754. .2097 of consensus"
24291. .24581
/note="AluSx repeat: matches 1. .300 of consensus"
24653. .24850
repeat_region /note="MIR repeat: matches 1. .200 of consensus"
26189. .26685
misc_feature /note="CpG island"
/evidence=not_experimental
29453. .30918
/note="CpG island"
/evidence=not_experimental
31084. .31156
repeat_region /note="MIR repeat: matches 65. .138 of consensus"
31618. .31876
repeat_region /note="AluB repeat: matches 29. .275 of consensus"
32767. .32830
repeat_region /note="MIR repeat: matches 76. .139 of consensus"
33050. .33178
/note="43 copies 3 mer tcc 72% conserved"
35112. .35201
misc_feature /note="CpG island"
/evidence=not_experimental
36551. .36604
/note="27 copies 2 mer ac 94% conserved"
37585. .38254
repeat_region /note="L1PA15 repeat: matches 5480. .6157 of consensus"
38265. .38561
repeat_region /note="AluSx repeat: matches 1. .300 of consensus"
39214. .39423
repeat_region /note="MIR repeat: matches 3. .219 of consensus"
39489. .39801
repeat_region /note="AluSx repeat: matches 1. .292 of consensus"
40050. .40189
repeat_region /note="MIR repeat: matches 1. .144 of consensus"
41057. .41290
/note="MIR repeat: matches 8. .240 of consensus"
41236. .41460
repeat_region /note="U2 repeat: matches 2569. .2730 of consensus"
41506. .41944
repeat_region /note="U2 repeat: matches 1916. .2416 of consensus"
42388. .42698
repeat_region /note="U2 repeat: matches 1448. .1779 of consensus"
44193. .44579
/note="THE1C repeat: matches 1. .371 of consensus"
44600. .44956
repeat_region /note="L1R16A repeat: matches 90. .445 of consensus"
45240. .45300
/note="MER8A repeat: matches 37. .97 of consensus"
45798. .45909
repeat_region /note="L1R41 repeat: matches 90. .192 of consensus"
46826. .46871
/note="23 copies 2 mer gt 97% conserved"
50684. .50969
repeat_region /note="AluSg repeat: matches 1. .287 of consensus"
50960. .51291
repeat_region /note="AluSx repeat: matches 1. .312 of consensus"
52222. .52519
repeat_region /note="AluSx repeat: matches 1. .300 of consensus"
54065. .54260
repeat_region /note="L1M4 repeat: matches 3865. .4055 of consensus"
54261. .54432
/note="FAM repeat: matches 2. .167 of consensus"
54433. .54629

Db 160849 GTTACCTGCTGTCCTCCCTGATGACCGGTGACCTCTTGAAGGCTGATAGGGGTGGTTT 160790
QY 554 GTTGAAGAGGACTTGTGCTGGCCCTTGTGTGAGAGCAGGATATTTGGAGGGATCTGTGTG 613
Db 160789 GTTGAAGAGGACTTGTGCTGGCCCTTGTGTGAGAGCAGGATATTTGGAGGGATCTGTGTG 160730
QY 614 GTGCTTGAAGGTATGATCAGAGAGGGAGCAGAGGTGTGTCTTCCCTTGTGTGTAG 673
Db 160729 GTGCTTGAAGGTATGATCAGAGAGGGAGCAGAGGTGTGTCTTCCCTTGTGTGTAG 160670
QY 674 CGTAGAGCAGAGGAGAGCTTATGTCAGCATTTCCAAAGTGTGGGTGGTCCGTTGGTTC 733
Db 160669 CGTAGAGCAGAGGAGAGCTTATGTCAGCATTTCCAAAGTGTGGGTGGTCCGTTGGTTC 160610
QY 734 CCGAGATATCTTTAGGTGTGTGAGGGGCTTGCAATTAAGTGCCACAAATCAGAGCAGAGAA 793
Db 160609 CCGAGAG-ACCTTTAGGTGTGTGAGGGGCTTGCAATTAAGTGCCACAAATCAGAGCAGAGAA 160551
QY 794 GCGATGCCCTTCCCAATTTCTCAATCCTTTTATGCGAGAGAGATCTCAGCTGATGCCA 853
Db 160550 GCGATGCCCTTCCCAATTTCTCAATCCTTTTATGCGAGAGAGATCTCAGCTGATGCCA 160491
QY 854 ACATGTTCCGATGCTGTGAGAGACATGCCAGCTCTCTCTGCTTGAAGGAGAGACTT 913
Db 160490 ACATGTTCCGATGCTGTGAGAGACATGCCAGCTCTCTCTGCTTGAAGGAGAGACTT 160431
QY 914 GGGCTTGAAGGAGGAG 973
Db 160430 GGGCTTGAAGGAGGAG 160371
QY 974 ATATTTTATTTGCTTATCTTTATTTGTTAGAGACAGTGTGATGAGCATTTCTATTATTG 1033
Db 160370 ATATTTTATTTGCTTATCTTTATTTGTTAGAGACAGTGTGATGAGCATTTCTATTATTG 160311
QY 1034 ACCTTTTCAATTAATGATTTAAGTMAAATAAATAAATAAATAAATAAATAAATAA 1081
Db 160310 ACCTTTTCAATTAATGATTTAAGTMAAATAAATAAATAAATAAATAAATAAATAA 160263

RESULT 6
LOCUS AR413227 483 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 864 from patent US 6639063.
ACCESSION AR413227
VERSION AR413227.1 GI:40168337
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 483)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 864 28-OCT-2003;
FEATURES
source 1..483
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 39.7%; Score 473; DB 6; Length 483;
Best Local Similarity 100.0%; Pred. No. 7.9e-108;
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 131 TTTGTGAGACCCATCAGGAGCCACAGAGGCTGGAGAGATGTGGATCTTGGCCAA 150
QY 181 TGGGAATGTGTGAGAGACAGACACCCCGAGTGAAGACACATACCAGACCAAGAG 240
Db 191 TGGGAATGTGTGAGAGACAGACACCCCGAGTGAAGACACATACCAGACCAAGAG 250
QY 241 TAGCATTCCTGACAGAGAGCTTCTTCAATAGAGGAGCATGTGTCTCCCGAGGGGTCTTG 300
Db 251 TAGCATTCCTGACAGAGAGCTTCTTCAATAGAGGAGCATGTGTCTCCCGAGGGGTCTTG 310
QY 301 CCCCCGACAGCAGAGCAGAGTGCAGAGCTGTGTCTCAATGCCCTTCAATGACCT 360
Db 311 CCCCCGACAGCAGAGCAGAGTGCAGAGCTGTGTCTCAATGCCCTTCAATGACCT 370
QY 361 CAACGAGAGCTGTGAGACATGAGGCTTCCGACAGTGCATCTCGCAACCATGCTGTGA 420
Db 371 CAACGAGAGCTGTGAGACATGAGGCTTCCGACAGTGCATCTCGCAACCATGCTGTGA 430
QY 421 GCGGATGACTCCATCT 473
Db 431 GCGGATGACTCCATCT 483

RESULT 7
LOCUS BD108780 483 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD108780
VERSION BD108780.1 GI:23203598
KEYWORDS JP 2002010789-A/857;
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 483)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 857 15-JAN-2002;
COMMENT GENSET CORP
OS Homo sapiens (human)
PN JP 2002010789-A/857
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280899
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..483
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 39.7%; Score 473; DB 6; Length 483;
Best Local Similarity 100.0%; Pred. No. 7.9e-108;
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 TTCTGTGAGCCCATCAGGGAACCAAGCGCTGGGAGATGTGCGANTTTGGCCA 180
Db 131 TTCTGTGAGCCCATCAGGGAACCAAGCGCTGGGAGATGTGCGANTTTGGCCA 190
QY 181 TGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 191 TGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250
QY 241 TAGCATTTCTCGACAGAGCTTCTTCAATAGGGGAGAGAGAGAGAGAGAGAGAG 300
Db 251 TAGCATTTCTCGACAGAGCTTCTTCAATAGGGGAGAGAGAGAGAGAGAGAGAG 310
QY 301 CCCCCGACAG 360
Db 311 CCCCCGACAG 370
QY 361 CACCGGAGCTGTGTGAACATGGGCTTTCGAGTGGAGAGAGAGAGAGAGAGAG 420
Db 371 CACCGGAGCTGTGTGAACATGGGCTTTCGAGTGGAGAGAGAGAGAGAGAGAG 430
QY 421 GCCGCTGACCTCCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 473
Db 431 GCCGCTGACCTCCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 483

RESULT 8
AX113082 550 bp DNA linear PAT 01-MAY-2001
LOCUS Sequence 5 from Patent WO0127276.
DEFINITION AX113082
ACCESSION AX113082
VERSION AX113082.1 GI:13939514
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Harlocker, S.L., Dillon, D.C. and Xu, J.
AUTHORS DNA sequences from breast tumor and uses thereof
JOURNAL Patent: WO 0127276-A 5 19-APR-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..550
Location/Qualifiers
1..550
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..550
/note="n = A,T,C or G"

ORIGIN
misc_feature
Query Match 38.0%; Score 452.6; DB 6; Length 550;
Best Local Similarity 95.4%; Pred. No. 1.1e-102;
Matches 519; Conservative 0; Mismatches 17; Indels 8; Gaps 5;

QY 47 CGCGCTGACAGACAGACTTACTAGCGGTGGGTCACTGTGAACATCACTGACTGC 106
Db 1 CGCGCTGACAGACAGACTTACTAGCGGTGGGTCACTGTGAACATCACTGACTGC 60
QY 107 AACCTTCCTCAATTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 166
Db 61 AACCTTCCTCAATTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 167 CGGATCTTGGCCAAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 226
Db 121 CGGATCTTGGCCAAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 227 CAGCCACCAAGAGAGTATCTCTCGACAGAGCTTCTCAATAGGGGAGAGAGAGAG 286
Db 181 CAGCCACCAAGAGAGTATCTCTCGACAGAGCTTCTCAATAGGGGAGAGAGAGAG 240
QY 287 CAGAGGGGCTCTGAG 346
Db 241 CAGAGGGGCTCTGAG 300

QY 347 CCTTCAATGACTCAACCGGACAGTGTGAACATGGGCTTTCCGAGATGATCTCGGC 406
Db 301 CCTTCAATGACTCAACCGGACAGTGTGAACATGGGCTTTCCGAGATGATCTCGGC 360
QY 407 AACCATGC-TGTGAGCCGGTGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 465
Db 361 AACCATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
QY 466 TGTGTGAG 524
Db 420 TGTGTGAG 479
QY 525 GACCTCT--GAGGAGCTGATAGAG--GTGGCTTTGTGAAGAGAGAGAGAGAGAG 579
Db 480 GACCTTTTGAAG 539
QY 580 TGTG 583
Db 540 TGGG 543

RESULT 9
AX431087 643 bp DNA linear PAT 28-JUN-2002
LOCUS Sequence 30 from Patent WO240535.
DEFINITION AX431087
ACCESSION AX431087
VERSION AX431087.1 GI:21656064
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Reipon, H., Liu, C., Salceda, S., Sun, Y., Caffery, R. and
AUTHORS Macina, R.A.
TITLE Compositions and methods relating to ovary specific genes and
JOURNAL proteins
Patent: WO 0240535-A 30 23-MAY-2002;
DIADEMUS INC (US)
FEATURES
source 1..643
Location/Qualifiers
1..643
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 36.3%; Score 432.8; DB 6; Length 643;
Best Local Similarity 96.1%; Pred. No. 1.1e-97;
Matches 519; Conservative 0; Mismatches 12; Indels 9; Gaps 7;

QY 662 CTTTGTGTTAAGCGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 104 CTTTGTGTTAAGCGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 163
QY 721 GGTCCGTTGTTCCCGAGATCTTTTAGTGTGATGGGCTCGATTAGTGGAC-AAA 779
Db 164 GGTCCGTTGTTCCCGAGATCTTTTAGTGTGATGGGCTCGATTAGTGGACAAA 223
QY 780 ATCAGAGCAAGAAAGCGATGCTTT-CCCATTTCTTCAATCTTTT--ATGCCGAGAG 836
Db 224 ATCAGAGCAAGAAAGCGATGCTTTCCCATTTCTTCAATCTTTTATGGCCGAGAG 283
QY 837 ATTCAGCTGATGCAACATGTTCCGATGCTGTGAGAGAGAGAGAGAGAGAGAGAG 896
Db 284 ATTCAGCTGATGCAACATGTTCCGATGCTGTGAGAGAGAGAGAGAGAGAGAGAG 343
QY 897 CCTAG 954
Db 344 CCTAG 403
QY 955 CTGTTTCTTTTATGATATAT-TTATTTGGTACTTTATTTTATGAGAGAGAGTGT 1013

Db 404 TGTATTTTGTATATGATATATTTCTATATGCTACTTATTTATGATAGACAAGTGTA 463
QY 1014 GTGGCATCTATTATTAT-TGTACCTTTTCATATAATAGATTAAATGAAAAA 1072
Db 464 GTGGCATCTATTATTATGATAGCTTTTCATATAATAGATTAAAGCAAAAAA 523
QY 1073 AAAAAAAAAAAGCTGTGGGGGTTTACCGGGCCAAAGGGGGCCCGGGGGGGAATCT 1132
Db 524 AAAAAAAAAAAGCTGTGGGGGTTTACCGGGCCAAAGGGGGCCCGGGGGGGAATCT 583
QY 1133 GGTTCCTCCGCCCCCAATTTCCCAATTTTGTAGCAAAATGAAGAGACACAGAAACC 1192
Db 584 GGTTCCTCCGCCCCCAATTTCCCAATTTTGTAGCAAAATGAAGAGACACAGAAACC 643

RESULT 10
AR415427 519 bp DNA linear PAT 18-DEC-2003
LOCUS AR415427
DEFINITION Sequence 3064 from patent US 6639063.
ACCESSION AR415427
VERSION AR415427.1 GI:40170537
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 519)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITL E EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 3064 28-OCT-2003;
FEATURES
source 1..519
location/Qualifiers
1..519
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 32.9%; Score 391.8; DB 6; Length 519;
Best Local Similarity 98.8%; Pred. No. 2.1e-87;
Matches 404; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
QY 54 GCAGACACAGCATCTACTGAGCGTGGTCACTCTGTGAACATCACTGACTGCAAGCTC 113
Db 112 GAATACACACATCTACTGAGCGTGGTCACTCTGTGAACATCACTGACTGCAAGCTC 171
QY 114 CCTCAATTTTGTGTGACGCCCATCAAGGACCCACAGCGCTGGAGGAGATGTGCGATCT 173
Db 172 CCTCAATTTTGTGTGACGCCCATCAAGGACCCACAGCGCTGGAGGAGATGTGCGATCT 231
QY 174 TGGCCAAATGGGAAATGTGTGACGAGACGAGCCCGAGTGAAGGACCATCAACCGCCAC 233
Db 232 TGGCCAAATGGGAAATGTGTGACGAGACGAGCCCGAGTGAAGGACCATCAACCGCCAC 291
QY 234 CAAGAGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGCCATGTGTCTCCCGAGGG 293
Db 292 CAAGAGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGCCATGTGTCTCCCGAGGG 350
QY 294 GTCTGTGCCCCCGCAGACAGAGGAGGTGCGAGGCTGGGTGTGCTCTGCTCATGCTTCA 353
Db 351 GTCTGTGCCCCCGCAGACAGAGGAGGTGCGAGGCTGGGTGTGCTCTGCTCATGCTTCA 410
QY 354 ATGACCTCAACCGGACGCTGTGAACATGGGCTTTCCGAGTGGACATCTCGGCAACATG 413
Db 411 ATGACCTCAACCGGACGCTGTGAACATGGGCTTTCCGAGTGGACATCTCGGCAACATG 470
QY 414 CTGTGAGCGGTGACCTCACTCTGCTCTTCTCTGCTCATGATGCT 462
Db 471 CTGTGAGCGGTGACCTCACTCTGCTCTTCTCTGCTCATGATGCT 519

RESULT 11
BD110980 519 bp DNA linear PAT 18-SEP-2002
LOCUS BD110980
DEFINITION EST and encoded human protein.

VERSION BD110980.1 GI:23205798
KEYWORDS JP 2002010789-A/3057.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 519)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITL E EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 3057 15-JAN-2002;
COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/3057
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
CC Key Location/Qualifiers
FT Key Location/Qualifiers
FT CDS 93..431.
1..519
location/Qualifiers
1..519
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 32.9%; Score 391.8; DB 6; Length 519;
Best Local Similarity 98.8%; Pred. No. 2.1e-87;
Matches 404; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
QY 54 GCAGACACAGCATCTACTGAGCGTGGTCACTCTGTGAACATCACTGACTGCAAGCTC 113
Db 112 GAATACACACATCTACTGAGCGTGGTCACTCTGTGAACATCACTGACTGCAAGCTC 171
QY 114 CCTCAATTTTGTGTGACGCCCATCAAGGACCCACAGCGCTGGAGGAGATGTGCGATCT 173
Db 172 CCTCAATTTTGTGTGACGCCCATCAAGGACCCACAGCGCTGGAGGAGATGTGCGATCT 231
QY 174 TGGCCAAATGGGAAATGTGTGACGAGACGAGCCCGAGTGAAGGACCATCAACCGCCAC 233
Db 232 TGGCCAAATGGGAAATGTGTGACGAGACGAGCCCGAGTGAAGGACCATCAACCGCCAC 291
QY 234 CAAGAGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGCCATGTGTCTCCCGAGGG 293
Db 292 CAAGAGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGCCATGTGTCTCCCGAGGG 350
QY 294 GTCTGTGCCCCCGCAGACAGAGGAGGTGCGAGGCTGGGTGTGCTCTGCTCATGCTTCA 353
Db 351 GTCTGTGCCCCCGCAGACAGAGGAGGTGCGAGGCTGGGTGTGCTCTGCTCATGCTTCA 410
QY 354 ATGACCTCAACCGGACGCTGTGAACATGGGCTTTCCGAGTGGACATCTCGGCAACATG 413
Db 411 ATGACCTCAACCGGACGCTGTGAACATGGGCTTTCCGAGTGGACATCTCGGCAACATG 470
QY 414 CTGTGAGCGGTGACCTCACTCTGCTCTTCTCTGCTCATGATGCT 462
Db 471 CTGTGAGCGGTGACCTCACTCTGCTCTTCTCTGCTCATGATGCT 519

RESULT 12
G22561 480 bp DNA linear STS 31-MAY-1996
LOCUS G22561/c
DEFINITION human STS WI-12709, sequence tagged site.
ACCESSION G22561
VERSION G22561.1 GI:1342887
KEYWORDS STS; STS sequence; primer; sequence tagged site.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAC plates: 36 Row: h Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
Location/Qualifiers

FEATURES
source

1. 1538
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MGC:28201 IMAGE:3989515"
/tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."
/clone_1kb="MGI CGAP_Lm29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1538
/gene="2010107G23R1k"
/db_xref="LocusID:69894"
/db_xref="MGI:1917144"
486..848
/codon_start=1
/product="2010107G23R1k protein"
/protein_id="AAH24943.1"
/db_xref="GI:19354289"
/db_xref="LocusID:69894"
/translation="MYRIILNGEIVDDDDPRVTTTTHSSSSQGFNRRHGAPRG
PGRRODAGARLQAQSPFSDLRQLVKNVGFPMHIGNVVEVTSILLFILMLGV
RGLLVVLVTVLSHSR"

CDS

ORIGIN

Query Match 25.8%; Score 307.4; DB 10; Length 1538;
Best Local Similarity 72.6%; Pred. No. 4,7e-66;
Matches 484; Conservative 0; Mismatches 166; Indels 17; Gaps 6;

119 ATTTCTGTGACGCCATCAGGACCCACAGCG-CTTGAGAGATGGTGGATCTTGGC 177
443 ATTACAGATCTGACAGCAGAAATTCACGCGTCCAGGAGAAATGATGATCTTGGC 502
178 CAATGGGGAATCGTGAGGACGACGACCCCGAGTGGAGACCACTACCGACCAAG 237
503 CAATGGGGAATCGTGAGGATGATGACCAAGATGAGACCAACCCAG---CACAG 559
238 AGTAGATTCCTTCAGACAGCTTTTCAATAGGGGCAATGTGCTTCCCGAGGGGCTC 297
560 AAGTAGAGCTCTCAGACAGGGCTTTTCAACAGAGGCAAGCGGCACTTCAGAGGGGCC 619
298 TGGCCCCCGCAGAGAGGAGGAGGCTGAGGCTGGGTGCTCAATGCCCTTCAATA 357
620 TGAACCCCGCAGAGAGGAGGAGGCTGAGGCTGGGTGCTCAATGCCCTTCAATA 679
358 CCTCAACCCGAGAGTGTGAAATAGGGCTTTCGCGAGTGCATCTCGGCAACCATCTGT 417
680 CTTGAACCCGAGAGTGTGAAATAGGGCTTTCGCGAGTGCATCTCGGCAACCATCTGT 739
418 GGAACCGGTGACCTTCATCTGCTCTCTTCCTGCTCATATGCTGTGGTCTGTGGCT 477
740 GGAACCGGTGACCTTCATCTGCTCTCTTCCTGCTCATATGCTGTGGTCTGTGGCT 799
478 CCTCTGGTGGCTCTTCTACCTGATGCTCCACCTGAGTCAAGGCTGACCTTGAAGG- 536
800 CTTGCTTGTGGGCTGTGCTCACTGCTGTCTCACTGAGTCAAGGCTGACCTTGAAGG- 859
537 CTGATAGGGGCTGTGCTGTGAGAGGA-----CTTGTGGGCTTGTGTGAGAGAGG 591
860 CTGATAGGGGCTGTGCTGTGAGAGGA-----CTTGTGGGCTTGTGTGAGAGAGG 651

Db 920 CACAGGGGAGGAGAGCTTGGGTGGGCTCAGAGGTGTGATGGGAGTACAGACCGGCTTGT 979
QY 652 TGTGTTTCCCTTTTGTGTTAAGCGTGAGGACAGAGACGTTAGTCCAGATTTCCAAA 711
Db 980 AGATTTTCTCCCGT-----GCAGAGGACAGAGGAGCTCTCTGGGCAAGATTCGGA 1034
QY 712 GTGGGTGGGTCCGTTGTTCCGAGATCTTTAGGTGATAGGCGCTGCATTAACT 771
Db 1035 --AAAGAGAGTCTGTGTTCTGTGAAATGTTTTAGGGGAGATAGGCGCTTGGCTAAA 1092
QY 772 GGACAAA 778
Db 1093 AGACAAA 1099
RESULT 14
AC116233/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-95117, *** SEQUENCING IN PROGRESS
AC116233
275924 bp DNA linear HTG 22-SEP-2002
Rattus norvegicus
AC116233.7 GI:23101097
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 275924)
REFERENCE
AUTHORS
Munzy, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, D.,
Allen, C., Allen, H., Alebrooks, S., Amth, A., Anguiano, D.,
Aryalbechi, V., Ayagi, A., Ayodeji, M., Bacc, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, N., Barnstead, M., Benamed, F.,
Blewalo, K., Blair, C., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasari, H., Catter, Z.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Deigado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Dval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Felle, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Gatta, M.,
Gebregioris, E., Geer, K., Gill, R., Gaddy, M., Guerra, M., Guervara, W.,
Gumarane, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Huix, S., Hume, J., Idebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowals, C., Kraft, C.L., Lebow, H., Levan, D., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheva, L., Louisse, H., Lorado, R.J., Lu, X., Ma, J.,
Maneswari, M., Mahindaratne, M., Mamoud, M., Malloy, K., Mangum, A.,
Mangum, B., Marple, P., Martin, K., Martin, R., Martinez, E.,
Mawhinley, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, N., Morris, K., Morris, S., Muniz, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwokelam, O., Okunnu, G., Olarnunsgoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Plasmoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzio, B., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaj, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Soelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Sytek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmami, K.,
Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, D.,

Query Match	24.7%;	Score 294;	DB 2;	Length 240657;
Best Local Similarity	51.6%;	Need No. 4	As 63;	

```

best local similarity 1.00; freq.no. 4.4e+02;
Matches 499; Conservative 0; Mismatches 181; Indels 17; Gaps 8;

```

QY	123	CTGTGTAGGCGCAATCAAGGGGCCCAAGGCGCT--GGAGAGATGGTGGGATCTTGGCCAACT	181
Db	51661	CAGGTCCAGGCCACGACGAAATTCAACAGCTCTTGGGATTAATGGTGAAGATCTTGGCTAAAT	51602
QY	182	GGGGAAATCGTGAGAGACGACGACCCCGAGTAGAGCACTAACCGAGCACCAGAGGT	241
Db	51601	GGGGAGATCGTTACGATGATGATACCCCGAGTAGAGCACTAACCCCA--ACAGAGAGT	51545
QY	242	AGGATTCCTGAGACGAGCTTCTCAATATGGGGGCAATGGTGTCTCCCAAGGGGCTCT--GG	300
Db	51544	AGCAGCACTCCGCAAGGAGCTTTTCAACAAGGGGCAAGTGCACCTTCCAGGGGCGCTTGGG	51485
QY	301	CCCCCGCAGACAGGACGAGGATGCCAGGCTGGGTGCTGTCACTATCCCTTCAATAGCT	360
Db	51484	ACCCNCGACAGACAGGACAGGATGCCAGCTGGGTGCTGCCAATCTCTTTGATGACCT	51425
QY	361	CAACCCGACACGTGTGAACAATGGGGCTTTCGGAGTAGGCAATCTCGGACAATGCTGTGA	420
Db	51424	GAACCCGACACGTGTGAACAATGGGGCTTTCGGAGTAGGCAATCTTTGGAAACACTGTGTGA	51365
QY	421	GCCGAGTACCTGCATCTGTGCTCTCTCTCTCAATGATGCTTGATGCTGTGGCCCTCT	480
Db	51364	GCCGTGACCTGCATCTGTGCTCTCTCTCTCAATGATGCTTGATGCTGTGGGCTCTCT	51305
QY	481	CTGTGTGGGCTTGTCTACTCTGTGTTCACACTGAATCAAGGATGACTCTTGAAGGCTGA	540
Db	51304	GCTTGTGGGCTTGTCTACTCTGTGTTCACACTGAATCAAGGATGACTCTTGAAGGCTGA	51245
QY	541	TAAAGGATGGGTTTGTGTAGAGGAGCTTGTCTGGGCTTGTGGTGTAGAGCAGCATATTTGG	600
Db	51244	TAAAGCAGGAGATGTTCATGATAGATGACATGGG--CTTCAATATGATATCAACGACAGGGG	51186
QY	601	AGGAGATCTGTGTGTGCTTGAAGATATGATCAAGAGGGGACCAAGGTGTGTGTTTCC	660
Db	51185	AGGAGAGCTTGTGTGTGCTTGAAGATATGATCAAGAGGGGACCGTGTGTGTGATCTTCC	51129
QY	661	CCTTTGTGTTAAACGTGAGGACAGGAGGAGAGCTTATGTCCAGACTTTCAAAAGTGTGGG	720
Db	51128	CC-----CGCTCAGAGAGGACAGGAGGAGCTCTTGTGGCCAGCAATGCTGAATAAATAG--A	51076
QY	721	GGTCCGTGGTGTCCCGAGATACTTTTATGGTGTATGGGGCTGTGCATTAATGGGACCAAA	780
Db	51075	GTCCTCATGTTTCTGTGAATGTTTTTATGGGAGCATATGGGCTTCATTACAAAGCAC--AAG	51017
QY	781	TCAGAGCAAGAAAGCGATGCCCTCTCCCAATCTCTCA	817
Db	51016	CGAGATGAAGAAAGCTTATGATGTTGTATTTTCCCA	50980

Search completed: August 22, 2004, 02:37:28
Job time : 4696 secs

Job time : 4696 secs

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 16:20:58 ; Search time 514 Seconds
(without alignments)

9851.852 Million cell updates/sec

Title: US-10-001-885-31

Sequence: 1192 1 aaaccgcccgaagcgcgcg.....gaagagacacacgaacc 1192

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_23Jan04:*

- 1: Geneseq1808:*
- 2: Geneseq19908:*
- 3: Geneseq20008:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2003as:*
- 8: Geneseq2003bs:*
- 9: Geneseq2003cs:*
- 10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1192	100.0	1192	6	ABN87820 Human ova
2	1061.4	89.0	1108	6	ABL90057 Human pol
3	1012	84.9	1049	2	AAV59558 Human sec
4	1012	84.9	1049	6	ABN873545 Human CDN
5	1012	84.9	1049	8	ACD82688 Human sequ
6	562	47.1	562	6	ABZ11212 Human pol
7	455.4	38.2	1165	5	AA564820 DNA encod
8	452.6	38.0	550	4	AA82505 Human bre
9	434.6	36.5	587	5	AA564821 DNA encod
10	432.8	36.3	643	6	ABN87819 Human ova
11	372.2	31.2	517	7	ABX74673 Human CDN
12	363.4	30.5	365	7	ABX92251 Human ova
13	363.2	30.5	612	7	ABX92250 Human ova
14	159	13.3	348	5	AA564819 DNA encod
15	75.6	6.3	461	6	ABT07555 Human bre
16	75.6	6.3	633	6	ABT07560 Human bre
17	75	6.3	986	6	ABT13409 Breast sp
18	71.2	6.0	420	6	ABT07648 Human bre
19	71.2	6.0	2269	6	ABT07649 Human bre
20	69.6	5.8	577	6	ABX92021 Lung spec
21	67.8	5.7	355	7	ABX92192 Human ova
22	67.4	5.7	999	6	ABN87265 Breast sp
23	67	5.6	668	6	ABT04062 Human ova

24	65.2	5.5	737	7	ABT13319 Breast sp
25	64.8	5.4	766	6	ABT07679 Human bre
26	64.4	5.4	1400	6	ABT07680 Human bre
27	64	5.4	666	6	ABN87812 Human ova
28	63.6	5.3	536	6	ABT07561 Human bre
29	63.6	5.3	626	6	ABN87873 Human ova
30	63.6	5.3	657	7	ABX92246 Human ova
31	63.6	5.3	1206	6	ABT07562 Human bre
32	62.8	5.3	652	6	ABN87291 Breast sp
33	62.2	5.2	716	6	ABT04069 Human ova
34	62.2	5.2	879	7	ABT13356 Breast sp
35	61.8	5.2	601	6	ABV83649 Human bre
36	61	5.1	422	4	AA115924 Human bre
37	60.4	5.1	800	7	ABX92232 Human ova
38	60.4	5.1	1545	6	ABX91992 Human ova
39	59.8	5.0	1623	6	ABD39107 Lung spec
40	59.6	5.0	464	5	ABV03810 Human pro
41	59.2	5.0	816	6	AB567268 Breast sp
42	58.4	4.9	510	6	ABN87806 Human ova
43	57.8	4.8	1495	7	ABX92216 Human ova
44	57.8	4.8	2910	6	ABN87829 Human ova
45	57.4	4.8	665	7	ABX92249 Human ova

ALIGNMENTS

RESULT 1	ABN87820	ABN87820 standard; cDNA, 1192 BP.
ID	ABN87820	
AC	ABN87820;	
DT	12-AUG-2002	(first entry)
DE	Human ovary specific nucleic acid SEQ ID NO:31.	
XX	Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;	
KW	ovary specific gene; OSNA; ovarian cancer; immune response; metastasis;	
KW	chromosome 10; gene; ss..	
XX		
OS	Homo sapiens.	
XX		
PN	WO200240535-A2.	
PD	23-MAY-2002.	
XX		
XX	20-NOV-2001; 2001WO-US045011.	
PF		
XX	20-NOV-2000; 2000US-0252061P.	
PR	27-NOV-2000; 2000US-0253257P.	
XX		
PA	(DIAD-) DIADEXUS INC.	
XX		
PI	Salceda S, Macina RA, Recipon H, Caffery R, Sun Y, Liu C;	
XX		
DR	WPI; 2002-471617/50.	
XX		
PT	New ovary specific genes and proteins, useful as a vaccine for treating	
PT	patients with ovarian cancer, or for diagnosing and monitoring the	
PT	presence and metastases of ovarian cancer in a patient.	
XX		
PS	Claim 1; Page 173-174; 260pp; English.	
XX		
CC	ABN87790 to ABN87882 represent human ovary specific nucleic acid (OSNA)	
CC	sequences, and ABN879297 to ABN89370 represent ovary specific protein	
CC	(OSP) sequences from the present invention. OSNA and OSP sequences have	
CC	cytostatic activity, and can be used in vaccine production and gene	
CC	therapy. An antibody that specifically binds to an OSP can be used for	
CC	treating a patient with ovarian cancer, particularly for inducing an	
CC	immune response against the ovarian cancer cell expressing the OSNA or	
CC	OSP. The OSNAs and OSPs can also be used for diagnosing and monitoring	
CC	the presence and metastases of ovarian cancer in a patient	

Oy		103	CCCTTCATTAATGATTAGTAAAAA	1077
Dd		1021	CCTTTCAATAATAGATTAAAGTAAAAA	1063
RESULT 3				
AAV59558				
ID	AAV59558	standard; DNA; 1049 BP.		
XX				
AC	AAV59558;			
XX				
DT	06-JAN-1999	(first entry)		
XX				
DE	Human secreted protein gene 48 clone HFCAT14.			
KW	Human; secreted protein; fusion protein; gene therapy; protein therapy;			
KM	diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;			
KW	developmental abnormality; foetal deficiency; blood; allergy; renal; ds;			
KM	immune system; aschma; lymphocytic disease; brain; hepatic; lymphoma;			
KW	inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;			
KM	cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;			
KW	osteoporosis; arthritis; testis; lung; thyroiditis; thyroid digestion;			
KX	endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.			
XX				
OS	Homo sapiens.			
XX				
PN	W09839448-A2.			
XX				
PD	11-SEP-1998.			
XX				
PF	06-MAR-1998;	98WC-US004493.		
XX				
PR	07-MAR-1997;	97US-0038621P.		
PR	07-MAR-1997;	97US-0040161P.		
PR	07-MAR-1997;	97US-0040162P.		
PR	07-MAR-1997;	97US-0040163P.		
PR	07-MAR-1997;	97US-0040333P.		
PR	07-MAR-1997;	97US-0040334P.		
PR	07-MAR-1997;	97US-0040336P.		
PR	07-MAR-1997;	97US-0040626P.		
PR	11-APR-1997;	97US-0043311P.		
PR	11-APR-1997;	97US-0043312P.		
PR	11-APR-1997;	97US-0043313P.		
PR	11-APR-1997;	97US-0043314P.		
PR	11-APR-1997;	97US-0043315P.		
PR	11-APR-1997;	97US-0043368P.		
PR	11-APR-1997;	97US-0043369P.		
PR	11-APR-1997;	97US-0043376P.		
PR	11-APR-1997;	97US-0043378P.		
PR	11-APR-1997;	97US-0043580P.		
PR	11-APR-1997;	97US-0043669P.		
PR	11-APR-1997;	97US-0043670P.		
PR	11-APR-1997;	97US-0043671P.		
PR	11-APR-1997;	97US-0043672P.		
PR	11-APR-1997;	97US-0043674P.		
PR	23-MAY-1997;	97US-0047492P.		
PR	23-MAY-1997;	97US-0047500P.		
PR	23-MAY-1997;	97US-0047501P.		
PR	23-MAY-1997;	97US-0047502P.		
PR	23-MAY-1997;	97US-0047503P.		
PR	23-MAY-1997;	97US-0047581P.		
PR	23-MAY-1997;	97US-0047582P.		
PR	23-MAY-1997;	97US-0047583P.		
PR	23-MAY-1997;	97US-0047584P.		
PR	23-MAY-1997;	97US-0047585P.		
PR	23-MAY-1997;	97US-0047586P.		
PR	23-MAY-1997;	97US-0047587P.		
PR	23-MAY-1997;	97US-0047588P.		
PR	23-MAY-1997;	97US-0047589P.		
PR	23-MAY-1997;	97US-0047590P.		
PR	23-MAY-1997;	97US-0047592P.		
PR	23-MAY-1997;	97US-0047593P.		

PR 23-MAY-1997; 97US-0047594P.
PR 23-MAY-1997; 97US-0047595P.
PR 23-MAY-1997; 97US-0047596P.
PR 23-MAY-1997; 97US-0047597P.
PR 23-MAY-1997; 97US-0047598P.
PR 23-MAY-1997; 97US-0047599P.
PR 23-MAY-1997; 97US-0047600P.
PR 23-MAY-1997; 97US-0047601P.
PR 23-MAY-1997; 97US-0047612P.
PR 23-MAY-1997; 97US-0047613P.
PR 23-MAY-1997; 97US-0047614P.
PR 23-MAY-1997; 97US-0047615P.
PR 23-MAY-1997; 97US-0047617P.
PR 23-MAY-1997; 97US-0047618P.
PR 23-MAY-1997; 97US-0047632P.
PR 23-MAY-1997; 97US-0047633P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048974P.
PR 13-JUN-1997; 97US-0049610P.
PR 08-JUL-1997; 97US-0051926P.
PR 16-JUL-1997; 97US-0052874P.
PR 18-AUG-1997; 97US-0055724P.
PR 22-AUG-1997; 97US-0056830P.
PR 22-AUG-1997; 97US-0056831P.
PR 22-AUG-1997; 97US-0056832P.
PR 22-AUG-1997; 97US-0056836P.
PR 22-AUG-1997; 97US-0056837P.
PR 22-AUG-1997; 97US-0056862P.
PR 22-AUG-1997; 97US-0056862P.
PR 22-AUG-1997; 97US-0056845P.
PR 22-AUG-1997; 97US-0056864P.
PR 22-AUG-1997; 97US-0056872P.
PR 22-AUG-1997; 97US-0056874P.
PR 22-AUG-1997; 97US-0056875P.
PR 22-AUG-1997; 97US-0056876P.
PR 22-AUG-1997; 97US-0056877P.
PR 22-AUG-1997; 97US-0056878P.
PR 22-AUG-1997; 97US-0056879P.
PR 22-AUG-1997; 97US-0056880P.
PR 22-AUG-1997; 97US-0056881P.
PR 22-AUG-1997; 97US-0056882P.
PR 22-AUG-1997; 97US-0056884P.
PR 22-AUG-1997; 97US-0056886P.
PR 22-AUG-1997; 97US-0056887P.
PR 22-AUG-1997; 97US-0056888P.
PR 22-AUG-1997; 97US-0056889P.
PR 22-AUG-1997; 97US-0056892P.
PR 22-AUG-1997; 97US-0056893P.
PR 22-AUG-1997; 97US-0056894P.
PR 22-AUG-1997; 97US-0056903P.
PR 22-AUG-1997; 97US-0056908P.
PR 22-AUG-1997; 97US-0056909P.
PR 22-AUG-1997; 97US-0056910P.
PR 22-AUG-1997; 97US-0056911P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057659P.
PR 05-SEP-1997; 97US-0057761P.
PR 12-SEP-1997; 97US-0058785P.
PR 02-OCT-1997; 97US-0061060P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC,
PI Bednarek DP, Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM,
PI Ferrite AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA,
PI Moore PA, Shi Y, Lafleur DM, Li Y, Zeng Z, Kyaw H;
XX
XX WPI: 1998-506364/43.
DR P-PSDB; AAW74778.

PT New isolated human genes and the secreted polypeptide(s) they encode -

PT disorders, immune diseases, inflammation or blood disorders.

XX Claim 1, Page 280; 721pp; English.

XX
XX This sequence represents a nucleic acid molecule designated Gene 48 from
CC the human cDNA clone HFCAT74 (deposited as clone ATCC 97899 and ATCC
CC 209045) which encodes a secreted human protein. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. AAV59502) for increasing the stability of the fused
CC protein as compared to the human protein only. The invention relates to
CC 186 novel genes and their fragments (nucleic acid sequences: AAV59511-
CC V59812; amino acid sequences AAW74731-W75026) which are useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. Also, pathological conditions can be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 186 polynucleotides, based on
CC which tissues they are most highly expressed in (see AAV59511 for
CC described uses).

XX
SQ Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

Query Match 84.9%; Score 1012; DB 2; Length 1049;
Best Local Similarity 99.3%; Pred. No. 1,28-222;
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 46 TCGCGCCCTGCAGACACGACATCTACTACGCGTGGGTCACTCTGTGAACATCATGACTG 105

DB 1 TCGCGCCCTGCAGACACGACATCTACTACGCGTGGGTCACTCTGTGAACATCATGACTG 60

QY 106 CAAGCTCCCTCAATTCTGTGTGAGCCCATCAGGACCCACAGCGCTGGAGATGAT 165

DB 61 CAAGCTCCCTCAATTCTGTGTGAGCCCATCAGGACCCACAGCGCTGGAGATGAT 120

QY 166 GCGATCTTGGCCAAATGGGGAAATCGTGCAGAGAGAGACCCCGAGTGGAGACCACTAC 225

DB 121 GCGATCTTGGCCAAATGGGGAAATCGTGCAGAGAGAGACCCCGAGTGGAGACCACTAC 180

QY 226 CGAGCCACAGAGATGACATCTCTGCAGACAGAGCTTCTCAATAGGGGCCATGGTGTCC 285

DB 181 CGAGCCACAGAGATGACATCTCTGCAGACAGAGCTTCTCAATAGGGGCCATGGTGTCC 240

QY 286 CGCAGGGGGCTCTGCCCCCGCCAGCAGCAGAGAGAGTGCAGGCTGGGTGTCTCATGC 345

DB 241 CGCAGGGGGCTCTGCCCCCGCCAGCAGCAGAGAGAGTGCAGGCTGGGTGTCTCATGC 300

QY 346 CCCCTTCAATGACCTCAACCGGCGCTGTGTGAATGAGGCTTTCCGAGTGGCATCTCCG 405

DB 301 CCCCTTCAATGACCTCAACCGGCGCTGTGTGAATGAGGCTTTCCGAGTGGCATCTCCG 360

QY 406 CAACCATGCTGTGAGCGCGGTGACCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 465

DB 361 CAACCATGCTGTGAGCGCGGTGACCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420

QY 466 TGTTCGTGGGCT 525

DB 421 TGTTCGTGGGCT 480

QY 526 ACCTCTGAGGGCTGATAGAGGGGTGGGTTTGTGAGAGGAACTTGTGGGCTTGTGTGAG 585

DB 481 ACCTCTGAGGGCTGATAGAGGGGTGGGTTTGTGAGAGGAACTTGTGGGCTTGTGTGAG 540

QY 586 AGCAGGCAATTTTGGAGGGGATCTGTGTGAGGCTTGAAGATGATCAGAGGGGACCA 645

DB 541 AGCAGGCAATTTTGGAGGGGATCTGTGTGAGGCTTGAAGATGATCAGAGGGGACCA 600

QY 646 CAGGTGTGTGTTTCCCTTGTGTGAAGCGTGAAGCAGAGGAGACCTTATGTC--AGCA 703

DB 601 CAGGTGTGTGTTTCCCTTGTGTGAAGCGTGAAGCAGAGGAGACCTTATGTC--AGCA 660

QY 704 TTTCCAAAGTGTGGTGGTCCCTTGTGTGTTCCGAGATCACTTTAAGGTGATGGGCGCTG 763

720

QY 764 CATTAGTGCACAAATTCAGAGCAGAGAAAGCATGCCCTTCCCAATTCCTCAATCCCTT 823
Db 721 CATTAGTGCACAAATTCAGAGCAGAGAAAGCATGCCCTTCCCAATTCCTCAATCCCTT 780
QY 824 TTATGCCGAGAAATCTCAGCTGGATGCCAAATGTTCCGATGCTCTGGAGAACATGCC 883
Db 781 TTATGCCGAGAAATCTCAGCTGGATGCCAAATGTTCCGATGCTCTGGAGAACATGCC 840
QY 884 GACGCTCTCTGCTTCTAGGAGAGAGCATTTGGCTTAGGGCAGCTGAGAAAAATTCACA 943
Db 841 GACGCTCTCTGCTTCTAGGAGAGAGCATTTGGCTTAGGGCAGCTGAGAAAAATTCACA 900
QY 944 CTTTATTAGACCTGTTTTTTTAAATGCTAATTTTATGCTACTTATTTGTTTAA 1003
Db 901 CTTTATTAGACCTGTTTTTTTAAATGCTAATTTTATGCTACTTATTTGTTTAA 960
QY 1004 ACAAGTGTAGTGCATTTCTATTATTGACCTTTTCAATAAATAGATTAAAGTAAAA 1063
Db 961 ACAAGTGTAGTGCATTTCTATTATTGACCTTTTCAATAAATAGATTAAAGTAAAA 1020
QY 1064 AAAAAAAAAAAAAA 1077
Db 1021 AAAAAAAAAAAAAA 1034

RESULT 4
ABS73545
ID ABS73545 standard; cDNA; 1049 BP.
XX
AC ABS73545;
XX
DT 15-JAN-2003 (first entry)
XX
DE Human cDNA #1 for novel secreted protein gene 48.
XX
KW Human; ss; gene; secreted protein; autoimmune disease; chemotaxis;
KW rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;
KW liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis; sunburn;
KW nervous system disorders; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; tissue regeneration;
KW epithelial cell proliferation; organ transplantation; food additive;
KW preservative; nutritional.
XX
OS Homo sapiens.
XX
PN US6420526-B1.
PD 16-JUL-2002.
XX
PF 08-SEP-1998; 98US-00149476.
XX
PR 07-MAR-1997; 97US-0038621P.
PR 07-MAR-1997; 97US-0040161P.
PR 07-MAR-1997; 97US-0040162P.
PR 07-MAR-1997; 97US-0040163P.
PR 07-MAR-1997; 97US-0040333P.
PR 07-MAR-1997; 97US-0040334P.
PR 07-MAR-1997; 97US-0040336P.
PR 07-MAR-1997; 97US-0040626P.
PR 11-APR-1997; 97US-0043311P.
PR 11-APR-1997; 97US-0043312P.
PR 11-APR-1997; 97US-0043313P.
PR 11-APR-1997; 97US-0043314P.
PR 11-APR-1997; 97US-0043315P.
PR 11-APR-1997; 97US-0043568P.
PR 11-APR-1997; 97US-0043569P.
PR 11-APR-1997; 97US-0043576P.
PR 11-APR-1997; 97US-0043578P.
PR 11-APR-1997; 97US-0043580P.
PR 11-APR-1997; 97US-0043669P.
PR 11-APR-1997; 97US-0043670P.

PR 11-APR-1997; 97US-0043671P.
PR 11-APR-1997; 97US-0043672P.
PR 11-APR-1997; 97US-0043674P.
PR 23-MAY-1997; 97US-0044932P.
PR 23-MAY-1997; 97US-0047500P.
PR 23-MAY-1997; 97US-0047501P.
PR 23-MAY-1997; 97US-0047502P.
PR 23-MAY-1997; 97US-0047503P.
PR 23-MAY-1997; 97US-0047581P.
PR 23-MAY-1997; 97US-0047582P.
PR 23-MAY-1997; 97US-0047583P.
PR 23-MAY-1997; 97US-0047584P.
PR 23-MAY-1997; 97US-0047585P.
PR 23-MAY-1997; 97US-0047586P.
PR 23-MAY-1997; 97US-0047587P.
PR 23-MAY-1997; 97US-0047588P.
PR 23-MAY-1997; 97US-0047589P.
PR 23-MAY-1997; 97US-0047590P.
PR 23-MAY-1997; 97US-0047592P.
PR 23-MAY-1997; 97US-0047593P.
PR 23-MAY-1997; 97US-0047594P.
PR 23-MAY-1997; 97US-0047595P.
PR 23-MAY-1997; 97US-0047596P.
PR 23-MAY-1997; 97US-0047597P.
PR 23-MAY-1997; 97US-0047598P.
PR 23-MAY-1997; 97US-0047599P.
PR 23-MAY-1997; 97US-0047600P.
PR 23-MAY-1997; 97US-0047601P.
PR 23-MAY-1997; 97US-0047612P.
PR 23-MAY-1997; 97US-0047613P.
PR 23-MAY-1997; 97US-0047614P.
PR 23-MAY-1997; 97US-0047615P.
PR 23-MAY-1997; 97US-0047616P.
PR 23-MAY-1997; 97US-0047617P.
PR 23-MAY-1997; 97US-0047618P.
PR 23-MAY-1997; 97US-0047619P.
PR 06-JUN-1997; 97US-0048964P.
PR 13-JUN-1997; 97US-0048974P.
PR 08-JUL-1997; 97US-0049610P.
PR 16-JUL-1997; 97US-0051928P.
PR 18-AUG-1997; 97US-0052874P.
PR 22-AUG-1997; 97US-0055724P.
PR 22-AUG-1997; 97US-0056630P.
PR 22-AUG-1997; 97US-0056631P.
PR 22-AUG-1997; 97US-0056632P.
PR 22-AUG-1997; 97US-0056633P.
PR 22-AUG-1997; 97US-0056634P.
PR 22-AUG-1997; 97US-0056635P.
PR 22-AUG-1997; 97US-0056636P.
PR 22-AUG-1997; 97US-0056637P.
PR 22-AUG-1997; 97US-0056638P.
PR 22-AUG-1997; 97US-0056639P.
PR 22-AUG-1997; 97US-0056640P.
PR 22-AUG-1997; 97US-0056641P.
PR 22-AUG-1997; 97US-0056642P.
PR 22-AUG-1997; 97US-0056643P.
PR 22-AUG-1997; 97US-0056644P.
PR 22-AUG-1997; 97US-0056645P.
PR 22-AUG-1997; 97US-0056646P.
PR 22-AUG-1997; 97US-0056647P.
PR 22-AUG-1997; 97US-0056648P.
PR 22-AUG-1997; 97US-0056649P.
PR 22-AUG-1997; 97US-0056650P.
PR 22-AUG-1997; 97US-0056651P.
PR 22-AUG-1997; 97US-0056652P.
PR 22-AUG-1997; 97US-0056653P.
PR 22-AUG-1997; 97US-0056654P.
PR 22-AUG-1997; 97US-0056655P.
PR 22-AUG-1997; 97US-0056656P.
PR 22-AUG-1997; 97US-0056657P.
PR 22-AUG-1997; 97US-0056658P.
PR 22-AUG-1997; 97US-0056659P.
PR 22-AUG-1997; 97US-0056660P.
PR 22-AUG-1997; 97US-0056661P.
PR 22-AUG-1997; 97US-0056662P.
PR 22-AUG-1997; 97US-0056663P.
PR 22-AUG-1997; 97US-0056664P.
PR 22-AUG-1997; 97US-0056665P.
PR 22-AUG-1997; 97US-0056666P.
PR 22-AUG-1997; 97US-0056667P.
PR 22-AUG-1997; 97US-0056668P.
PR 22-AUG-1997; 97US-0056669P.
PR 22-AUG-1997; 97US-0056670P.
PR 22-AUG-1997; 97US-0056671P.
PR 22-AUG-1997; 97US-0056672P.
PR 22-AUG-1997; 97US-0056673P.
PR 22-AUG-1997; 97US-0056674P.
PR 22-AUG-1997; 97US-0056675P.
PR 22-AUG-1997; 97US-0056676P.
PR 22-AUG-1997; 97US-0056677P.
PR 22-AUG-1997; 97US-0056678P.
PR 22-AUG-1997; 97US-0056679P.
PR 22-AUG-1997; 97US-0056680P.
PR 22-AUG-1997; 97US-0056681P.
PR 22-AUG-1997; 97US-0056682P.
PR 22-AUG-1997; 97US-0056683P.
PR 22-AUG-1997; 97US-0056684P.
PR 22-AUG-1997; 97US-0056685P.
PR 22-AUG-1997; 97US-0056686P.
PR 22-AUG-1997; 97US-0056687P.
PR 22-AUG-1997; 97US-0056688P.
PR 22-AUG-1997; 97US-0056689P.
PR 22-AUG-1997; 97US-0056690P.
PR 22-AUG-1997; 97US-0056691P.
PR 22-AUG-1997; 97US-0056692P.
PR 22-AUG-1997; 97US-0056693P.
PR 22-AUG-1997; 97US-0056694P.
PR 22-AUG-1997; 97US-0056695P.
PR 22-AUG-1997; 97US-0056696P.
PR 22-AUG-1997; 97US-0056697P.
PR 22-AUG-1997; 97US-0056698P.
PR 22-AUG-1997; 97US-0056699P.

PR 22-AUG-1997; 97US-0056910P.
 PR 22-AUG-1997; 97US-0056911P.
 PR 05-SEP-1997; 97US-0057650P.
 PR 05-SEP-1997; 97US-0057659P.
 PR 05-SEP-1997; 97US-0057761P.
 PR 12-SEP-1997; 97US-0058785P.
 PR 02-OCT-1997; 97US-0061060P.
 PR 06-MAR-1998; 98MO-US004493.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC,
 PI Bednarik DR, Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM,
 PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA,
 PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;

XX WPI; 2002-634796/68.

DR P-PSDE; AB935227.

PT New isolated human secreted protein for diagnosing, preventing, treating
 PT or ameliorating medical conditions and used as a food additive or
 PT preservative.

PS Example 1; SEQ ID NO 58; 129pp; English.

XX The invention relates to an isolated protein that is one of 186 human
 CC secreted proteins, given in the specification, encoded by one of 309 cDNA
 CC sequences also given in the specification. The protein is used in a
 CC pharmaceutical composition used to prevent, treat or ameliorate a medical
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
 CC chickens or sheep. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
 CC infections caused by bacteria, viruses and fungi and ocular disorders
 CC e.g. corneal infection. The polypeptides can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional
 CC components. The present sequence represents a cDNA derived from a gene
 CC encoding one of the novel human secreted proteins of the invention. Note:
 CC This sequence did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=642052681

SO Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

Query Match 84.9%; Score 1012; DB 6; Length 1049;

Best Local Similarity 99.3%; Pred. No. 1.2e-222; Indels 2; Gaps 1;

Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

DB 1 TCGGCGCTGACAGACAGCATCTACAGCGTGGGTCACTCTGTGAACATCACTGACTG 105

QY 46 TCGGCGCTGACAGACAGCATCTACAGCGTGGGTCACTCTGTGAACATCACTGACTG 105

DB 106 CAAGCTTCCTCAATTTCTGTGTGAGCCCATCAAGGACCCAGCGCTGGAGAGATGGT 165

QY 61 CAAGCTTCCTCAATTTCTGTGTGAGCCCATCAAGGACCCAGCGCTGGAGAGATGGT 120

DB 121 GCGGATCTTGGCCAAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

QY 226 CCAAGCCCAAGAGAGATGATCTCTGACAGAGAGCTTTCAATGAGGAGAGAGAGAGAG 285

DB 181 CCAAGCCCAAGAGAGATGATCTCTGACAGAGAGCTTTCAATGAGGAGAGAGAGAGAG 240

QY 286 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 345

DB 241 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

QY 346 CCCCTCAATGACCTCAACCGGAGAGCTGTGAACATGAGGCTTTCCGAGTGCATCTCGG 405

DB 301 CCCCTCAATGACCTCAACCGGAGAGCTGTGAACATGAGGCTTTCCGAGTGCATCTCGG 360

QY 406 CAACCATGCTGTGAGAGCGGAGAGCTTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 465

DB 361 CAACCATGCTGTGAGAGCGGAGAGCTTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 420

QY 466 TGTTCGTGAGCT 525

DB 421 TGTTCGTGAGCT 480

QY 526 ACCTCTGAGGCTGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 585

DB 481 ACCTCTGAGGCTGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

QY 586 AGCAGGCAATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 645

DB 541 AGCAGGCAATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600

QY 646 CAGGTGTGTGTTTCCCTTGTGTGTTAAGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 703

DB 601 CAGGTGTGTGTTTCCCTTGTGTGTTAAGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 660

QY 704 TTTCGAAGT 763

DB 661 TTTCGAAGT 720

QY 764 CATTAGTGACAAATTCAGAGCAAGAAAGCAATGCTTCCCAATCTCTCAATCTT 823

DB 721 CATTAGTGACAAATTCAGAGCAAGAAAGCAATGCTTCCCAATCTCTCAATCTT 780

QY 824 TTATGCCAGAAATCTCAGTGTGATGCCAATGTCCTGAGGAGGAGGAGGAGGAGGAGGAGG 883

DB 781 TTATGCCAGAAATCTCAGTGTGATGCCAATGTCCTGAGGAGGAGGAGGAGGAGGAGGAGG 840

QY 884 GACGCTCTCTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 943

DB 841 GACGCTCTCTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900

QY 944 CTTTATTGACATGTTTGTGTTTAAATGATATTTTATTTGCTTATTTATTTATTTATTTAGG 1003

DB 901 CTTTATTGACATGTTTGTGTTTAAATGATATTTTATTTATTTATTTATTTATTTATTTAGG 960

QY 1004 ACAAGTGTAGTGGATCTATTATTGTGACCTTTCAATTAATTAATTAATTAATTAATTA 1063

DB 961 ACAAGTGTAGTGGATCTATTATTGTGACCTTTCAATTAATTAATTAATTAATTAATTA 1020

QY 1064 AAAAAAAAAAAAAA 1077

DB 1021 AAAAAAAAAAAAAA 1034

QY 1021 AAAAAAAAAAAAAA 1034

DB 1021 AAAAAAAAAAAAAA 1034

QY 1021 AAAAAAAAAAAAAA 1034

DB 1021 AAAAAAAAAAAAAA 1034

QY 1021 AAAAAAAAAAAAAA 1034

DB 1021 AAAAAAAAAAAAAA 1034

QY 1021 AAAAAAAAAAAAAA 1034

DB 1021 AAAAAAAAAAAAAA 1034

QY 1021 AAAAAAAAAAAAAA 1034

DB 1021 AAAAAAAAAAAAAA 1034

QY 1021 AAAAAAAAAAAAAA 1034

DB 1021 AAAAAAAAAAAAAA 1034

QY 1021 AAAAAAAAAAAAAA 1034

KM respiratory disorder; neurological disorder; Alzheimer's disease;
KM Parkinson's disease; inflammation; Crohn's disease; vulvovaginitis;
KM immunosuppressive; antibacterial; haemostatic; thrombolytic;
KM anticoagulant; neuroprotective; thyromimetic; antiallergic;
KM antileukemic; virucide; fungicide; anti-HIV; nephrotoxic; antiangiinal;
KM cerebroprotective; cardiant; nootropic; antiparkinsonian;
KM antiinflammatory; gene; ss.
XX Homo sapiens.
OS
PN US2003049618-A1.
XX
PD 13-MAR-2003.
XX
PF 16-MAR-2001; 2001US-00809391.
XX
XX 07-MAR-1997; 97US-0038621P.
PR 07-MAR-1997; 97US-0040162P.
PR 07-MAR-1997; 97US-0040163P.
PR 07-MAR-1997; 97US-0040333P.
PR 07-MAR-1997; 97US-0040334P.
PR 07-MAR-1997; 97US-0040336P.
PR 07-MAR-1997; 97US-0040626P.
PR 11-APR-1997; 97US-0043311P.
PR 11-APR-1997; 97US-0043312P.
PR 11-APR-1997; 97US-0043313P.
PR 11-APR-1997; 97US-0043314P.
PR 11-APR-1997; 97US-0043315P.
PR 11-APR-1997; 97US-0043568P.
PR 11-APR-1997; 97US-0043569P.
PR 11-APR-1997; 97US-0043576P.
PR 11-APR-1997; 97US-0043578P.
PR 11-APR-1997; 97US-0043580P.
PR 11-APR-1997; 97US-0043669P.
PR 11-APR-1997; 97US-0043670P.
PR 11-APR-1997; 97US-0043671P.
PR 11-APR-1997; 97US-0043672P.
PR 11-APR-1997; 97US-0043674P.
PR 23-MAY-1997; 97US-0047492P.
PR 23-MAY-1997; 97US-0047500P.
PR 23-MAY-1997; 97US-0047501P.
PR 23-MAY-1997; 97US-0047502P.
PR 23-MAY-1997; 97US-0047503P.
PR 23-MAY-1997; 97US-0047581P.
PR 23-MAY-1997; 97US-0047582P.
PR 23-MAY-1997; 97US-0047583P.
PR 23-MAY-1997; 97US-0047584P.
PR 23-MAY-1997; 97US-0047585P.
PR 23-MAY-1997; 97US-0047586P.
PR 23-MAY-1997; 97US-0047587P.
PR 23-MAY-1997; 97US-0047588P.
PR 23-MAY-1997; 97US-0047589P.
PR 23-MAY-1997; 97US-0047590P.
PR 23-MAY-1997; 97US-0047592P.
PR 23-MAY-1997; 97US-0047593P.
PR 23-MAY-1997; 97US-0047594P.
PR 23-MAY-1997; 97US-0047595P.
PR 23-MAY-1997; 97US-0047596P.
PR 23-MAY-1997; 97US-0047597P.
PR 23-MAY-1997; 97US-0047598P.
PR 23-MAY-1997; 97US-0047599P.
PR 23-MAY-1997; 97US-0047600P.
PR 23-MAY-1997; 97US-0047601P.
PR 23-MAY-1997; 97US-0047612P.
PR 23-MAY-1997; 97US-0047613P.
PR 23-MAY-1997; 97US-0047614P.
PR 23-MAY-1997; 97US-0047615P.
PR 23-MAY-1997; 97US-0047617P.
PR 23-MAY-1997; 97US-0047618P.
PR 23-MAY-1997; 97US-0047632P.
PR 23-MAY-1997; 97US-0047633P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048974P.

PR 13-JUN-1997; 97US-0049610P.
PR 08-JUL-1997; 97US-0051926P.
PR 16-JUL-1997; 97US-0052874P.
PR 18-AUG-1997; 97US-0055724P.
PR 22-AUG-1997; 97US-005630P.
PR 22-AUG-1997; 97US-005631P.
PR 22-AUG-1997; 97US-005632P.
PR 22-AUG-1997; 97US-005633P.
PR 22-AUG-1997; 97US-0056637P.
PR 22-AUG-1997; 97US-0056639P.
PR 22-AUG-1997; 97US-0056642P.
PR 22-AUG-1997; 97US-0056643P.
PR 22-AUG-1997; 97US-0056644P.
PR 22-AUG-1997; 97US-0056872P.
PR 22-AUG-1997; 97US-0056874P.
PR 22-AUG-1997; 97US-0056875P.
PR 22-AUG-1997; 97US-0056876P.
PR 22-AUG-1997; 97US-0056877P.
PR 22-AUG-1997; 97US-0056878P.
PR 22-AUG-1997; 97US-0056879P.
PR 22-AUG-1997; 97US-0056880P.
PR 22-AUG-1997; 97US-0056881P.
PR 22-AUG-1997; 97US-0056882P.
PR 22-AUG-1997; 97US-0056883P.
PR 22-AUG-1997; 97US-0056884P.
PR 22-AUG-1997; 97US-0056885P.
PR 22-AUG-1997; 97US-0056887P.
PR 22-AUG-1997; 97US-0056908P.
PR 22-AUG-1997; 97US-0056909P.
PR 22-AUG-1997; 97US-0056910P.
PR 22-AUG-1997; 97US-0056911P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057668P.
PR 05-SEP-1997; 97US-0057761P.
PR 12-SEP-1997; 97US-0058785P.
PR 09-OCT-1997; 97US-0061660P.
PR 06-MAR-1998; 98MO-US004493.
PR 08-SEP-1998; 98US-00149476.
PR 17-MAR-2000; 2000US-0190066P.
XX
XX (RUBE/) RUBEN S M.
PA (ROSE/) ROSEN C A.
PA (SOPP/) SOPPET D R.
PA (CART/) CARTER K C.
PA (BEDN/) BEDNARIK D P.
PA (ENDR/) ENDRESS G A.
PA (YUGG/) YU G.
PA (NIJU/) NI J.
PA (FENG/) FENG P.
PA (YOUN/) YOUNG P E.
PA (GREE/) GREENE J M.
PA (FERR/) FERRIE A M.
PA (DUNN/) DUNN D R.
PA (HUUJ/) HU J.
PA (FLOR/) FLORENCE K A.
PA (OLSE/) OLSEN H S.
PA (FISC/) FISCHER C L.
PA (EBNE/) EBNER R.
PA (BREW/) BREWER L A.
PA (MOOR/) MOORE P A.
PA (SHIY/) SHI Y.
PA (LAFU/) LAFLEUR D W.
PA (LIYU/) LI Y.
PA (ZENG/) ZENG Z.
PA (KYAW/) KYAW H.
XX
PI Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP;

PI Endress GA, Yu G, Ni J, Peng P, Young PE, Greene JM, Ferrie AM,
PI Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Bener R,
PI Brewer LA, Moore PA, Shi Y, Tafleur DW, Li Y, Zeng Z, Kyaw H;
XX
DR WPI : 2003-521800/49.
DR P-PsDB; AB034421.

PT New genetic and its encoded prostate cancer antigen proteins, useful for
 PT preventing, treating, ameliorating or diagnosing e.g. prostate cancers,
 PT thymic hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral
 PT ischemia.
 XX
 PS
 PS Claim 4; SEQ ID NO 58; 260pp; English.

The present invention relates to the isolation of novel human secreted proteins and the polynucleotide sequences encoding them. The invention also discloses vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The polypeptide and polynucleotide sequences for the secreted proteins are useful for preventing, treating, ameliorating or diagnosing medical conditions such as hyperproliferative disorders (e.g., leukemia or breast cancers), wounds, reproductive disorders, blood-related disorders (e.g., haemophilia or thrombocytopaenia), immunodeficiencies (e.g., Wiskott-Aldrich syndrome or thymic hypoplasia), autoimmune disorders (e.g., graft-versus-host disease), multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g., asthma), viral or bacterial or fungal infections (e.g., AIDS or sepsis), renal disorders (e.g., kidney failure), cardiovascular disorders (e.g., angina pectoris, cerebral ischaemia or congenital heart defects), respiratory disorders, neurological disorders (e.g., Alzheimer's disease or Parkinson's disease), and inflammations (e.g., Crohn's disease). The polynucleotide or polypeptide may also be used as vaccine adjuvants. ACD82641-ACD82950 encode human secreted proteins or their fragments. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov/?seqIdentify.html

Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

Query Match	Score	ID	DB	Length
84.98%	1012	8	1049	

Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy	46	TTGGCGCTGCAACAACAACATCTA	CTCAGAGGTGGGTCACTCTGTGAACATCACTGAC	CTG	105
Db	1	TTGGCGCTGCAACAACAAGATCTTACT	CAGCGTGGTACTCTGTGAACATCACTGAC	CTG	60
Qy	106	CAAGCTTCCTCAATTTCTTGCTGAG	CCCATCAGAGACCCACAGCGCTTGGAGAGATG	T	165
Db	61	CAAGCTTCCTCAATTTCTTGCTGAG	CCCATCAGAGACCCACAGCGCTTGGAGAGATG	T	120
Qy	166	GGCGAATCTTGGCCATATGGGAAAT	ATGTGTGACAGAGAGAGAGACCCCGAGTGAAGACCACTAC		225
Db	121	GGCGAATCTTGGCCATATGGGAAAT	ATGTGTGACAGAGAGAGAGACCCCGAGTGAAGACCACTAC		180
Qy	226	CCAGCCACCAAGGATAGCATTTCT	CCGACAGAGCTTCTTCAATAGAGGACCATAGTGTCTC		285
Db	181	CCAGCCACCAAGGATAGCATTTCT	CCGACAGAGCTTCTTCAATAGAGGACCATAGTGTCTC		240
Qy	286	CCGAGGGGGTCTGTGGCCCCCGG	CCAGCAGCAGGTCGAGGCTGTGGTCTGCTAC	GT	345
Db	241	CCGAGGGGGTCTGTGGCCCCCGG	CCAGCAGCAGGTCGAGGCTGTGGTCTGCTAC	GT	300
Qy	346	CCCTTTCAATGACCTCAACCGG	CAGCTGTGAATAATGGGCTTTTCGCAATGGGACTCTGG		405
Db	301	CCCTTTCAATGACCTCAACCGG	CAGCTGTGAATAATGGGCTTTTCGCAATGGGACTCTGG		360
Qy	406	CAACCAATGTGTGAGACCGGTGA	CACTCCATCTGCTCTTTCCTCATGATGATGCTTGG		465
Db	361	CAACCAATGTGTGAGACCGGTGA	CACTCCATCTGCTCTTTCCTCATGATGATGCTTGG		420
Qy	466	TGTTGTGGCTCTTCTGTGGTGGC	CTTGTCTACCTGTGTGCTCCACTGATGACAGCGTG		525

QY	526	ACCTCTAGAGGCTATATGAGGGATGGGATTGTGTGAGAGGGACTTGTCTGGGCTTGGTGTGAG	583
Db	481	ACCTCTAGAGGCTATATGAGGGATGGGATTGTGTGAGAGGGACTTGTCTGGGCTTGGTGTGAG	540
QY	586	AGCAGGCATATTTGAGAGGGGATCTGCTGCTGTCCTTGAAGGTATATCAGAGAGGGACCA	645
Db	541	AGCAGGCATATTTGAGAGGGGATCTGCTGCTGTCCTTGAAGGTATATCAGAGAGGGACCA	600
QY	646	CAGGTGTGTGTTCCCTTTGTGTTAAGCGTGAAGGCAGAGGAGACGTTATGCC - AGCA	703
Db	601	CAGGTGTGTGTTCCCTTTGTGTTAAGCGTGAAGGCAGAGGAGACGTTATGCCAGCAT	660
QY	704	TTTCCAAAGTGTGGGTGGTCCGTTGGTTCCCGAGATCTTTAGTGTGATATGGGGCTG	763
Db	661	TTTCCAAAGTGTGGGTGGTCCGTTGGTTCCCGAGATCTTTAGTGTGATATGGGGCTG	720
QY	764	CATTATGTCGACAAATCAGAGCAAAAAGCAGTGCCTTCCCAATTCCTCAATCCTT	823
Db	721	CATTATGTCGACAAATCAGAGCAAAAAGCAGTGCCTTCCCAATTCCTCAATCCTT	780
QY	824	TTATGCCGAGAGATCTCAGCTGATGCCCAATGTTCCATGCTCTGTGAAACATAGCC	883
Db	781	TTATGCCGAGAGATCTCAGCTGATGCCCAATGTTCCATGCTCTGTGAAACATAGCC	840
QY	884	GAGCTCTCCTGCTGCTGAGGAGCAGGACTTGAGGCTTGAAGGAGGAGAAAAATTCAGA	943
Db	841	GAGCTCTCCTGCTGCTGAGGAGCAGGACTTGAGGCTTGAAGGAGGAGAAAAATTCAGA	900
QY	944	CTTTTATGACATGTGTTTGTTTTAATGTATATTTTATGCTTATGTTAGG	1003
Db	901	CTTTTATGACATGTGTTTGTTTTAATGTATATTTTATGCTTATGTTAGG	960
QY	1004	ACAAGCTGATGTCATCTTATATATGTGACCTTCAATTAATATGATTAAATAAAA	1063
Db	961	ACAAGCTGATGTCATCTTATATATGTGACCTTCAATTAATATGATTAAATAAAA	1020
QY	1064	AAAAAAAAAAAAAA 1077	
Db	1021	AAAAAAAAAAAAAA 1034	
RESULT 6			
ABZ11212			
ID	ABZ11212	standard; cDNA; 562 BP.	
XX	AC	ABZ11212;	
XX	20-JAN-2003	(first entry)	
DE	Human polynucleotide SEQ ID NO 94.		
XX	Human; genome mapping; gene therapy; food supplement; virus; fungus;		
XX	cell-proliferative disorder; neurodegenerative disease; bacterial;		
XX	Parkinson's disease; Alzheimer's disease; autoimmune disease;		
XX	multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;		
XX	arthritis; cytostatic; immunomodulator; nocotropic; neuroprotective;		
XX	antiparkinsonian; antidiabetic; immunosuppressive; dermatological;		
XX	haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;		
XX	antiarthritic; gene; ss.		
OS	Homo sapiens.		
XX	WO200270539-A2.		
XX	12-SEP-2002.		
XX	05-MAR-2002; 2002WO-US005095.		
XX	05-MAR-2001; 2001US-00799451.		
XX	/usrct \ \usrct \usr		


```
Db 308 AATGACAGTGCAGCCATCAGGAGCCCAAGCCCTGGAGAGATGTGCGGATCTTGG 367
Oy 177 CCAATGGGGAATGTGTGACAGGAC-GACGACCCCGAGTGAAGACCACTAACCCAGCCACCA 235
Db 368 CCAATGGGGAATGTGTGACAGGACGACGACCCCGAGTGAAGACCACTAACCCAGCCACCA 427
Oy 236 AGAGTGAATCTCTGACAGAGC-TTCTTAAATAGAGGAGCCATGTGTCTCCCAAGGAGG 294
Db 428 AGAGTGAATCTCTGACAGAGCCTTCTTAAAGAGGAGCCATGTGTCTCCCAAGGAGG 487
Oy 295 TCC-TGGCCCCCGGACAGAGAGS---CAGGTGCAAGGCTGGGTGCTGCTAG-TCCCCC 349
Db 488 TCCGTGCCCCCGGACAGAGAGS---CAGGTGCTGGGTGCTGCTGCTGCTGCTGCTGCTG 547
Oy 350 TTCAATGACC-TCAACCGGAGAGCTGTG-TAACATGGGCTTTCCGAGT--GGCATCTGG 405
Db 548 TTCAATGACCTTAAACCGGAGAGCTGTGTAACAATGGGCTTTCCGAGTGGCATCTTCCG 607
Oy 406 CAACCATGCTGTGAGCGCGGTGACCTGATCC-----TGCTCTCTTCTCCGCTCATGATG 460
Db 608 CAACCATGCTGTGAGCGCGGTGTAAGATCACTGCTGCTCTCTCTCTCTGCTCATGATG 667
Oy 461 CTGTGTGTGTGAGCTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
Db 668 CTGTGTGTGTGAGCTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 727
Oy 520 GCGGTGACCTTGAAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
Db 728 GCGGTGACCTTGAAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 787
Oy 580 TGTGAGAGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 634
Db 788 TGTGAGAGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 847
Oy 635 AGAGGAGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 692
Db 848 AGAGGAGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 907
Oy 693 TTACTCGAGATTTCCAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 752
Db 908 TTACTCGAGATTTCCAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 967
Oy 753 TATGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 812
Db 968 TATGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998
Oy 813 TCTCATCTCTTTATGCCGAGAGAGATCTGAGCTGATGCCAATGTTCGATGCTGTG 872
Db 999 ----- 998
Oy 873 GAAGACATGCCAGCTCTCTCTGCTGAGGAGAGAGCTTGGGCTTAGGAGAGAGAGAGAGAG 932
Db 999 -----GAGCAGAGACTTGGGCTTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1028
Oy 933 AAAATTCAGACTTTTATGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 992
Db 1029 AAAATTCAGACTTTTATGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1088
Oy 993 TATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1052
Db 1089 TATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1148
Oy 1053 TTAAGTAAAAA 1069
Db 1149 TTAAGTAAAAA 1165
```

```
RESULT 8
AAF82505
ID AAF82505 standard; cDNA; 550 BP.
XX AAF82505;
```

```
DT 18-JUN-2001 (first entry)
XX
XX Human breast tumour protein cDNA 15964.
DE Human; breast cancer; tumour; cytosolic; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX WO200127276-A2.
XX
XX 19-APR-2001.
XX
XX 11-OCT-2000; 2000WO-US028255.
XX
XX 12-OCT-1999; 99US-00417031.
XX
XX (CORI-) CORIXA CORP.
XX
XX Harlocker SL, Dillon DC, Xu J;
XX
XX WPI, 2001-273773/28.
XX
XX New polypeptides encoded by polynucleotide sequences over-expressed in
XX breast tumor tissue are useful to detect, monitor and treat breast
XX cancer.
XX
XX Claim 13; Page 50; 52pp; English.
XX
XX The present sequence encodes a breast tumour-associated protein. It was
XX shown to have at least two-fold overexpression in breast tumour tissue.
XX The invention relates to an isolated polypeptide comprising at least an
XX immunogenic portion of a breast tumor-specific protein, or its variant
XX that retains the ability to react with antigen-specific antisera. The
XX breast tumour polynucleotides, polypeptides and antibodies are useful for
XX inhibiting development of breast cancer. The polynucleotides may be used
XX to design primers and probes for detecting and monitoring breast cancer
XX
XX Sequence 550 BP; 102 A; 163 C; 164 G; 118 T; 0 U; 3 Other;
XX
XX Query Match 38.0%; Score 452.6; DB 4; Length 550;
XX Best Local Similarity 95.4%; Pred. No. 3.9e-94;
XX Matches 519; Conservative 0; Mismatches 17; Indels 8; Gaps 5;
Oy 47 GCGGCTGACAGACAGCATCTACTGACGCGTGTACCTGTGTGAACATCTGACTGC 106
Db 1 GCGGCTGACAGACAGCATCTACTGACGCGTGTACCTGTGTGAACATCTGACTGC 60
Oy 107 AAGCTCTCTCAATTTCTGTGTGAGCCCATGAGGAGCCAGGCGCTGGAGAGATGTTG 166
Db 61 AAGCTCTCTCAATTTCTGTGTGAGCCCATGAGGAGCCAGGCGCTGGAGAGATGTTG 120
Oy 167 CGGATCTTGGCCAAATGGGGAATGTGTGAGAGCAGACACCCCGAGTGAAGAGCAGTACC 226
Db 121 CGGATCTTGGCCAAATGGGGAATGTGTGAGAGCAGACACCCCGAGTGAAGAGCAGTACC 180
Oy 227 CAGCCACCAAGAGTATGATCTCTGACAGAGCTTTCTTCAATAGGGCCATGTGTCTCC 286
Db 181 CAGCCACCAAGAGTATGATCTCTGACAGAGCTTTCTTCAATAGGGCCATGTGTCTCC 240
Oy 287 CAGGGGGTCTGTGGCCCCCGGACAGAGAGCAGGTGACAGGCTGGGTGCTGCTGCTC 346
Db 241 CAGGGGGTCTGTGGCCCCCGGACAGAGAGCAGGTGACAGGCTGGGTGCTGCTGCTC 300
Oy 347 CCGTTCAATGACCTCAACCGGAGCTGTGTGAACATGGGCTTTCCGAGTGGCATCTGGC 406
Db 301 CCGTTCAATGACCTCAACCGGAGCTGTGTGAACATGGGCTTTCCGAGTGGCATCTGGC 360
Oy 407 AACCATGG-TGTGAGAGCGGTGACCTTCATCTGTCTCTGCTCATGATGCTTGG 465
Db 361 AACCATGG-TGTGAGAGCGGTGACCTTCATCTGTCTCTGCTCATGATGCTTGG 419
Oy 466 TGTGTGTGCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 524
```

Db 420 TGTGCTGAGCTTCTTGTGCTTGTCTACTGAGTCCCACTGATCAACGGCT 479
Qy 525 GACCTCT--GAGGCGTGATAGG--GTGGTTTGTGAGAGGACTTGTGGCCTTGG 579
Db 480 GACCTTTTGAAGGCGATGATGAGGATGTTGTGAAAAGAACTTGTGGCCTTGG 539
Qy 580 TGGG 583
Db 540 TGGG 543

RESULT 9
AAS64821
ID AAS64821 standard; cDNA; 587 BP.
AC AAS64821;
AT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #625.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS WO200175067-A2.
PN 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US008631.
PF 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
DR P-PsDB; ABG00634.
PI New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 625; 103bp; English.
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

Sequence 587 BP; 147 A; 106 C; 166 G; 168 T; 0 U; 0 Other;

Query Match 36.5%; Score 434.6; DB 5; Length 587;
Best Local Similarity 93.9%; Pred. No. 5.5e-90;
Matches 551; Conservative 0; Mismatches 24; Indels 12; Gaps 9;

Qy 506 TCCGACCTGAGTACAGCGGTGACCTCTGAGGGCTGATAGGGGTGGGTTGTGAGAGGAC 565
Db 1 TCCGACCTGAGTACAGCGGTGACCTCTGAGGGCTGATAGGGGTGGGTTGTGAGAGGAC 60
Qy 566 TTGCTGGGCTTGTGTGAGAGCAGGCAATATTGAGAGGGGATCTGTGCTCTTAAAG 625
Db 61 TTGCTGGGCTTGTGTGAGAGCAGGCAATATTGAGAGGGGATCTGTGCTCTTAAAG 120
Qy 626 -TATGATCAGAGAGGGG-ACCAAGAGTGTGTTTCCCTTGTGTTAAGGTTG---AGG 680
Db 121 TTATGATCAGAGAGGGGCCCCCAAGAGTGTGTTTCCCTTGTGTTAAGGTTG 180
Qy 681 CAGAGGAGAGCGTTAGTCC--AGCATTTCCAAAGTGTGAGTGGTCCGTTGGTT-CCCGA 737
Db 181 CAGAGGAGAGCGTTAGTCCAGCATTTCCAAAGTGTGAGTGGTCCGTTGGTTCCCGA 240
Qy 738 GATACCTTTAGTGGT-ATGGGGCTTGATTAAGTGGACAAATACAGACAAAGAACG 796
Db 241 GATACCTTTAGTGGTAAATGGGGCTTGATTAAGTGGACAAATACAGACAAAGAACG 300
Qy 797 ATGCCCTTCCCAATTCCTC-ATTCCTTTATGGCGAGAAAGATCTCAGCTGATG-CCAA 854
Db 301 ATGCCCTTCCCAATTCCTCATAATCCTTTATGGCGAGAAATTCAGTGAAGCCCA 360
Qy 855 CATGTTCCGATGCTGTGGAAGACATGCGCAGTCTCTGTGCTTAGGAGC-AGACTT 913
Db 361 CAGGTTCCGATGCTGTGGAAGACATGCGCAGTCTCTGTGCTTAGGAGCAGACTT 420
Qy 914 GGGCTTAGGAGGAGGAGGAAAAATTCAGACTTTTATAGCATGTTTGTGTTAATG 973
Db 421 GGGCTTAGGAGGAGGAGGAAAAATTCAGACTTTTATAGCATGTTTGTGTTAATG 480
Qy 974 ATATTTTATTTGGCTACTTATTTGTTTGAACAAGTGTAGTGCATTTATTTATG 1033
Db 481 ATATTTTATTTGGCTACTTATTTGTTTGAACAAGTGTAGTGCATTTATTTATG 540
Qy 1034 ACCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
Db 541 ACCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 587

RESULT 10
ABN87819
ID ABN87819 standard; cDNA; 643 BP.
AC ABN87819;
AT 12-AUG-2002 (first entry)
DE Human ovary specific nucleic acid SEQ ID NO:30.
KW Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;
KW ovary specific gene; OSG; ovarian cancer; immune response; metastasis;
XX gene; ss.
OS Homo sapiens.
XX WO200240535-A2.
PN 23-MAY-2002.
PD 20-NOV-2001; 2001WO-US045011.
PF 20-NOV-2000; 2000US-0252061P.
PR 27-NOV-2000; 2000US-0253257P.
XX (DIAD-) DIADEXUS INC.
PA Salceda S, Macina RA, Recipon H, Cafferty R, Sun Y, Liu C;
XX PI

Db 166 ----- 167
QY 852 CAACATGTCGATGCTGTGAAGACATGCCAGCTCTCTCTGCTAGGAGCAGAC 911
Db 166 -----GAGCAGGAC 158
QY 912 TTGGGCTTAGGGGAGGTGAAAAATTCAGACTTTTTCAGACTGTTTGTTCATTAAG 971
Db 157 TTGGGCTTAGGGGAGGTGAAAAATTCAGACTTTTTCAGACTGTTTGTTCATTAAG 98
QY 972 GTATATTTTATTTGGCTACTTTTATTTGTTAGACAAGTGTAGTGCATTCATTTATG 1031
Db 97 GTATATTTTATTTGGCTACTTTTATTTGTTAGACAAGTGTAGTGCATTCATTTATG 38
QY 1032 TGACCTTTTCATTAATGATTTAAGTAAAAAAA 1068
Db 37 TGACCTTTTCATTAATGATTTAAGTAAAAAATGAAA 1

RESULT 12
ABX92251
ID ABX92251 standard; cDNA; 365 BP.
XX
AC ABX92251;
XX
DT 08-MAY-2003 (first entry)
XX
DS Human ovarian specific nucleic acid DEX0310_65.
XX
KM Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;
KM non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;
KM gynaecological.
XX
OS Homo sapiens.
XX
PN WO200292785-A2.
XX
PD 21-NOV-2002.
XX
PF 13-FEB-2002; 2002WO-US022271.
XX
PR 13-FEB-2001; 2001US-0268290P.
PR 15-FEB-2001; 2001US-0268834P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferty R;
PI Sun Y, Liu C;
DR P-PSDB; AB061063.
XX
PT New isolated OSNA nucleic acid and encoded polypeptide, useful for
PT identifying, diagnosing, monitoring, staging, imaging and treating
PT ovarian cancer and non-cancerous diseases in ovarian tissues.
XX
PS Claim 1; Page 185-186; 224p; English.

The invention relates to a new isolated nucleic acid termed ovarian specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that encodes any of 53 fully defined protein sequences appearing as AB061018-AB061070 (termed ovarian specific proteins, OSP); (b) any of 76 fully defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a sequence having at least 60% sequence identity to the nucleic acid molecule of (a) or (b). Also included are a method for determining the presence of an ovary specific nucleic acid (OSNA) in a sample, a vector comprising an OSNA, a host cell comprising the vector, an isolated OSP polypeptide, an anti-OSP antibody or fragment, a method for determining the presence of an ovary specific protein in a sample and a vaccine comprising an OSP or OSNA. The methods and compositions of the present invention are useful for identifying, diagnosing, monitoring, staging, imaging and treating ovarian cancer and non-cancerous disease in ovary tissue. The present sequence is an OSNA of the invention

XX SQ Sequence 365 BP; 60 A; 117 C; 108 G; 80 T; 0 U; 0 Other;
Query Match 30.5%; Score 363.4; DB 7; Length 365;
Best Local Similarity 99.7%; Pred. No. 1,1e-73;
Matches 364; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 161 ATGTGCGATCTTGTGGCAATGGGAAATCGTGACAGACGACCCCGAGTGAAGAC 220
Db 1 ATGTGCGGATCTTGTGGCAATGGGAAATCGTGACAGATGACGACCCCGAGTGAAGAC 60
QY 221 ACTACCGACGACCAAGAGGTAGCATTCCTGACAGAGCTTCTCAATAGGGGCAATGAT 280
Db 61 ACTACCGACGACCAAGAGGTAGCATTCCTGACAGAGCTTCTCAATAGGGGCAATGAT 120
QY 281 GCTCCCGCAGGGGGTCTGTGGCCCCCGCAGACAGAGGAGTGCAGGCTGGTGTGCT 340
Db 121 GCTCCCGCAGGGGGTCTGTGGCCCCCGCAGACAGAGGAGTGCAGGCTGGTGTGCT 180
QY 341 CAGTCCCGCTTCATATGACCTCAACCGGAGCTGTGTGAACATGGGCTTCCGCAATGGAT 400
Db 181 CAGTCCCGCTTCATATGACCTCAACCGGAGCTGTGTGAACATGGGCTTCCGCAATGGAT 240
QY 401 CTCGCAACCATGCTGTGAGACCGGTGACCTCCATCCGTCTCTTCTGCTCATGATG 460
Db 241 CTCGCAACCATGCTGTGAGACCGGTGACCTCCATCCGTCTCTTCTGCTCATGATG 300
QY 461 CTGGGTGTTCTGTGGCTCTCTCTGTTGGCTTGTCTAAGCTGTGTCCACCTGATGAG 520
Db 301 CTGGGTGTTCTGTGGCTCTCTCTGTTGGCTTGTCTAAGCTGTGTCCACCTGATGAG 360
QY 521 CGGTG 525
Db 361 CGGTG 365

RESULT 13
ABX92250
ID ABX92250 standard; cDNA; 612 BP.
XX
AC ABX92250;
XX
DT 08-MAY-2003 (first entry)
XX
DS Human ovarian specific nucleic acid DEX0310_64.
XX
KM Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;
KM non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;
KM gynaecological.
XX
OS Homo sapiens.
XX
PN WO200292785-A2.
XX
PD 21-NOV-2002.
XX
PF 13-FEB-2002; 2002WO-US022271.
XX
PR 13-FEB-2001; 2001US-0268290P.
PR 15-FEB-2001; 2001US-0268834P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferty R;
PI Sun Y, Liu C;
DR P-PSDB; AB061062.
XX
PT New isolated OSNA nucleic acid and encoded polypeptide, useful for
PT identifying, diagnosing, monitoring, staging, imaging and treating
PT ovarian cancer and non-cancerous diseases in ovarian tissues.

DT 14-NOV-2002 (first entry)
XX
DE Human breast cancer associated coding sequence SEQ ID NO: 38.
XX
KM Human; breast specific gene; breast specific protein; breast cancer;
XX gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200264611-A1.
XX
PD 22-AUG-2002.
XX
PF 12-FEB-2002; 2002WO-US004197.
XX
PR 13-FEB-2001; 2001US-0268282P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferty R;
PI Sun Y, Liu C;
XX
DR WPI; 2002-657582/70.
XX
PT New breast specific nucleic acids and proteins, useful for identifying,
PT diagnosing, monitoring, staging, imaging, and treating breast cancer and
PT non-cancerous disease states in breast tissue, and in gene therapy.
XX
PS Claim 1, Page 179, 367BP, English.
XX
CC The present invention provides human breast specific coding sequences and
CC proteins. These can be used in the diagnosis and treatment of breast
CC cancer and non-cancerous diseases of the breast. The present sequence is
CC a coding sequence of the invention
XX
SQ Sequence 461 BP; 142 A; 84 C; 90 G; 145 T; 0 U; 0 Other;

Query Match 6.3%; Score 75.6; DB 6; Length 461;
Best Local Similarity 62.3%; Pred. No. 1.7e-07;
Matches 117; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 986 GCTACTTATTTGTTAGACAGTGTGCTGCTTCTATTATTGTGACCTTTCAATA 1045
DB 271 GCTTTTCTCTTTGTGCTCAAGTAGCTGCTTAAATTTGTGAAATATAAAC 330

QY 1046 AATAGATTATTAATAAAAAAAAAAAAAAAAAAAGTGTGGGGTTACCCGGG 1105
DB 331 TTTTGTATTAGAAAAAAGAAAAAAAAAAAAAAAAAGCTTGGGGAAACCCGGG 390

QY 1106 CCAAGGGGGCCCCCGGGGGGAAATGTGTTCCCGCCCAATTCCCCCAATTG 1165
DB 391 CCAAAAGCGGTGTCCGGGGGAATTGTTCTCCGTCCAAAATCCCCAAAAAATCG 450

QY 1166 AAAAA 1171
DB 451 AAAAA 456

Search completed: August 22, 2004, 01:19:02
Job time : 524 secs

Blank

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5629.f For
more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/htp://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF034CF12NP1&cluster=5629.f)
Feng Liang Email : liang@l1ftech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

QY 502 GGNGTCCCACTGAGTCAGCGGTACCTCTGAGGGCTGATAGGGGTGCGTTTGTGAGAG 561

Db 481 GGGTCTCCCACTGAGTCAGCGGTACTCTCTGAGGGCTGATAGGGGTGCG--TTTGTGAGAG 539

QY 562 GGACTTGTGCGGCTTGTGTGTGAGACAGACATATTGTAAGGGGATCTGTGTGTGCTTG 621

Db 540 GGACTTGTGCGGCTTGTGTGTGAGACAGACATTTTGTGAGGGGATCTGTGTGTGCTTG 599

QY 622 AAGGTATGATCAAGAGGGGAGACACAGGTGTGTGTTTCCCTTTGTGTATAAGCTGTAGGC 681

Db 600 AAGGTATGATCAAGAGGGGAGACACAGGTGTGTGTCTCCCTTGTGTATAAGGTGTAGGC 659

QY 682 AAGGGAGACGTTAGTCCAGCATTTCCCAAAGTGTGGGTGCGTGTGTTCCCGAGATA 741

Db 660 AAGGGAGACGTTAGTCC--AGCATTCCAAAGTGTGGGTGCGTGTG--TCCCGAGATA 716

QY 742 CTTTTAGTGTATGTGGGCTGTGACTTAAGTGCGCAAAATCAGAGCAAGAAAGCGATGCC 801

Db 717 CTTTACGGTGGAAATGGGG--CTGCAATTAGTGG--CAAATCTGAGCAAGAAAGCGAATGCC 773

QY 802 CTTCCCAATTCCTCTCAATCTTTTATGCCGAGAAAGATCTCAGCTGTAGTCCAAACATGTTC 861

Db 774 CTTCCCAATTCCTTAAACCTTTATATGCCGAGAAAAATTCAGGGGTATGTCCACATGTTC 833

QY 862 C 862

Db 834 C 834

ORIGIN

Query Match	Score	DB	Length
58.48;	696.2;	12;	708;

Matches 698; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

383 GGCTTCCGCA GTGCATCTCGGCA CCATGCTGTGAGCCGGTGACCTCCATCCTGCTC 442

701 GGCCTTCCGCA GTGGCATCTCGCA ACCATGCTGTGAGCCCGTGACCTCCATCCTGCTC 642

[illegible]

1. The first step in the process of creating a new product is to identify a market need. This involves conducting market research to understand what consumers want and what problems they are facing. Once a need is identified, the next step is to develop a concept that addresses that need. This is often done through brainstorming sessions with a team of designers and engineers. The concept is then refined through prototyping and testing, with feedback from potential users being used to make improvements. Finally, the product is launched into the market, and its success is monitored through sales data and customer feedback.

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

503 GTGTCACCTGACTCAGCGGTGACCTCTGAGGCTGATAGGGTGGGTTTGTTGAGAGG 562

581 GTGTCCACCTGAGTCAGCGGTGACCTCTGAGGCTGATAGGGTGGTTTGTGAGAGG 522

563 GACTTGCTGGCCCTTGGTGTGAGAGCAGGCATATTGGAGGGCATCTGGTGTGCCCTTGA 622

521 GACTTGGCTGGCCCTTGGTGTGAGAGCAGGCATATTGGAGGGGATCTGTTGTGTCCTTGA 462

523 AGGTAATGATCAGAGAGGAGCCACAGGTGTGTTCACCTTGTGTTAAAGCGTGAAGCA 682

402

100

[illegible]

743 TTTTACGTCGTATCGGGCCATTCAGTGGCACTAAATCAGAGCAAGAAAGCGATGCC 802

341 TTTTAGGTGATGGGCGTCATTAGTGGCACAATCAGAGCAAGAAAGCGATGCC 282

803 TTCCCAATTCTCAATCCTTTATGCGGAGAAGATCTCAGCTGGATGCCAACATGTTCC 862

281 TTCCAATTCTCTCAATCCTTTTATGCCGAGAATCTCAGCTGGATGCCCAACATGTTCC 222

863 GATGCCCTGTGGAAGACATGCCGACGTCCTCTGCTAGGAGCAGGACTTGGGCTTAGG 922

223 CATGCTCTGAGACCATCTCTCTCTGCTCTACGAGACGAGACTTGGGCTTAGG 162

THE UNIVERSITY OF CHICAGO

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

983 1GGCTACTTATGT1TAGGCAAG1GGAG1GGCA1C1A11A11G1GAC1111CA 10

101 TTGGCTACTTATGTTAGGACCAAGTGGTAGTGGCAATCTATTATTGTGACCTTTCA 42

Choroid"
/dev string="fetal and adult"
/lab host=MDH0B (Life Technologies) (T1 phage resistant)"
/clone lib="UT-E-EJ0"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polyniker; Site, 1: EcoR I; Site, 2: Not I;
UT-E-EJ0 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
RCAR/CNAGA, lens, CATTAGCGA, eye anterior segment,
FATGCGGATG, optic nerve, CCATTAGAGT, retina, CGCGG, Retina
Foveal and Macular, GTCC, RPE and Choroid, ACCTA. This
library was created for the Program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI).
TAG, LISB-UT-E-EJ0
TAG_SEQ=GTCC

TAG_SEQ=GTC"

QY 1043 ATAAATGATTAAATTAAGTAAAAA 1083
Db 41 ATAAATGATTAAATTAAGTAAAAA 1

RESULT 5
LOCUS B1521387 688 bp mRNA linear EST 29-AUG-2001
DEFINITION 603081177F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5220570 5', mRNA sequence.
ACCESSION B1521387
VERSION B1521387.1 GI:15346179
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 688)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM1554 row: 1 column: 19
High quality sequence stop: 688.
Location/Qualifiers

FEATURES
source 1..688
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5220570"
/lab_host="DH108"
/clone_lib="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb. insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 57.4%; Score 684.8; DB 12; Length 688;
Best Local Similarity 99.7%; Pred. No. 3.3e-86;
Matches 688; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 22 AGTGAAGCGCCGCGAGGCGCGAGTGGCGCTTGACAGACAGAGCATCTACAGCGTGGT 81
Db 1 AGTGAAGCGCCGCGAGGCGCGAGTGGCGCTTGACAGACAGAGCATCTACAGCGTGGT 60
QY 82 CACCTCTGGAATCATCATGCACTGCAAGCTCCCTCAATTTCTGTCGACGCCATCAGG 141
Db 61 CACCTCTGGAATCATCATGCACTGCAAGCTCCCTCAATTTCTGTCGACGCCATCAGG 120
QY 142 ACCCAGAGCGCTCGGAGAGATGGTGGCATCTTGGCCAAATGAGGAAATGTCAGAGCAG 201
Db 121 ACCCAGAGCGCTCGGAGAGATGGTGGCATCTTGGCCAAATGAGGAAATGTCAGAGAT 180
QY 202 CGACCCCGAGTGAAGACCACTACCGACCAACAAAGGTAGATCTCTGACAGAGCTT 261
Db 181 CGACCCCGAGTGAAGACCACTACCGACCAACAAAGGTAGATCTCTGACAGAGCTT 240
QY 262 CTTCATAGGGGCGCATGTGTCTCCCGAGGGGCTCTGGCCCCCGCCAGACAGAGGACAG 321

Db 241 CTTCATAGGGGCGCATGTGTCTCCCGAGGGGCTCTGGCCCCCGCCAGACAGAGGACAG 300
QY 322 TGCCAGGCTGGGTGCTGCTCAGTCCCTTCATATGACCTCAACGGGAGCTGTGAACAT 381
Db 301 TGCCAGGCTGGGTGCTGCTCAGTCCCTTCATATGACCTCAACGGGAGCTGTGAACAT 360
QY 382 GGGCTTTCCGAGTGGCATCTTCGGCAACCATGCTGTGAAGCCGGTACCTTCATCTGCT 441
Db 361 GGGCTTTCCGAGTGGCATCTTCGGCAACCATGCTGTGAAGCCGGTACCTTCATCTGCT 420
QY 442 CCGCTTCGTCATGATGCTTGGGTTCGAGGCTCCCTCGTGGTGGCTTGTCTACCT 501
Db 421 CCGCTTCGTCATGATGCTTGGGTTCGAGGCTCCCTCGTGGTGGCTTGTCTACCT 480
QY 502 GGTGTCCCACTGATGATGACGGGTGACCTTGAGGGCTGTGATAGGGGTGTTGAGAG 561
Db 481 GGTGTCCCACTGATGATGACGGGTGACCTTGAGGGCTGTGATAGGGGTGTTGAGAG 540
QY 562 GACCTGTGGGCTGCTGCTGTGAGAGAGGATATTTGAGAGGATCTGCTGCTGCT 621
Db 541 GACCTGTGGGCTGCTGCTGTGAGAGAGGATATTTGAGAGGATCTGCTGCTGCTGCT 600
QY 622 AAGGTATGATGAGAGAGGAGCAACAGGTGTGTGTTCCCTTTGTGTTAAGCGTGAAGC 681
Db 601 AAGGTATGATGAGAGAGGAGCAACAGGTGTGTGATCCCTTTGTGTTAAGCGTGAAGC 660
QY 682 AGAGGAGACCTTACTCCAGCATTTCCA 709
Db 661 AGAGGAGACCTTACTCCAGCATTTCCA 688

RESULT 6
LOCUS BM459665 937 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT 6417963 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534295
ACCESSION BM459665
VERSION BM459665.1 GI:18508705
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 937)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM1220 row: 1 column: 16
High quality sequence stop: 585.
Location/Qualifiers

FEATURES
source 1..937
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5534295"
/tissue_type="leiomyosarcoma"
/lab_host="DH108 (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

ORIGIN

Tue Aug 24 09:48:35 2004

us-10-001-885-31.rst

Page 6

Query Match	57.4%;	Score 684;	DB 12;	Length 937;
Best Local Similarity	97.1%;	Pred. No. 3.5e-86;		
Matches 707;	Conservative 0;	Mismatches 20;	Indels 1;	Gaps 1;

Qy	322	GCAGGAGTGCCTCAAGTCCCCCTTCAAGAAGCTCAACGGCACTGGTGAACATGGGCTT	387
Db	1	GCAGGAGTGCCTCAAGTCCCCCTTCAAGAAGCTCAACGGCACTGGTGAACATGGGCTT	60
Qy	388	TCCGCAATGAGCACTTCGGCAACATAGCTGTGAGACCCGGTGAAGCTCCATCTGCTCTCTT	447
Db	61	TCCGCAATGAGCACTTCGGCAACATAGCTGTGAGACCCGGTGAAGCTCCATCTGCTCTCTT	120
Qy	448	CCGCACTAGATGCTTGGTTCGAGGCTCTCCTCTGATTGGCTCTTGTCACTCCGATGTC	507
Db	121	CCGCACTAGATGCTTGGTTCGAGGCTCTCCTCTGATTGGCTCTTGTGTCACTCCGATGTC	180
Qy	508	CCAAGTGAATCAAGCGGTGAAGCTTCTGAGGCTGATTAAGGGTGGGTTGTTGAAGGGAATT	567
Db	181	CCAAGTGAATCAAGCGGTGAAGCTTCTGAGGCTGATTAAGGGTGGGTTGTTGAAGGGAATT	240
Qy	568	GCAGGAGCTTGGAGTGAAGGAGCAAGCATTTTGAAGAGGGAATCTGATGATGCTTGAAGTA	627
Db	241	GCAGGAGCTTGGAGTGAAGGAGCAAGCATTTTGAAGAGGGAATCTGATGATGCTTGAAGTA	300
Qy	628	TGATCAGAGAGGGAGCCACAGGTGTGTGTTCCCTTGTGTGAAGCGTGAAGGCAAGGG	687
Db	301	TGATCAGAGAGGGAGCCACAGGTGTGTGTTCCCTTGTGTGAAGCGTGAAGGCAAGGG	360
Qy	688	AGAGCTTAATCCAGCAATTCCAAAGTGGGGTGGTCCGTTGGTCCGAGATCACTTTTA	747
Db	361	AGAGCTTAATCCAGCAATTCCAAAGTGGGGTGGTCCGTTGGTCCGAGATCACTTTTA	420
Qy	748	GGTGTATGGGGCTTGCAATTAAGTGCCAAAAATCAGAGCAAGAAACGATGCCCTTCCC	807
Db	421	GGTGTATGGGGCTTGCAATTAAGTGCCAAAAATCAGAGCAAGAAACGATGCCCTTCCC	480
Qy	808	AATTCCTCAATCCCTTTTAATGCCAGGAAGATCTAGCTGAATGCCAATGTTCCGATGC	867
Db	481	AATTCCTCAATCCCTTTTAATGCCAGGAAGATCTAGCTGAATGCCAATGTTCCGATGC	540
Qy	868	CTGTGAGAGACATGCCAGATCTCCTCTGCTCTAGGAGCAAGGACTTGGGCTTAAAGGCAAG	927
Db	541	CTGTGAGAGACATGCCAGATCTCCTCTGCTCTAGGAGCAAGGACTTGGGCTTAAAGGCAAG	600
Qy	928	TGGAAAAAATTCAGACTTTTAACTTTTAAGCACTGTTTTGTATTAATGTAATTTTAAATGGC	987
Db	601	TGGAAAAAATTCAGACTTTTAACTTTTAACTGTTTTGTATTAATGGAATATTTTAAATGGC	660
Qy	988	TACTTAATGTTTAAAGACAAAGGTAGAGGCAATTCATTAATTTAGCACTTTTCAATA	1046
Db	661	TACTTAATGTTTAAAGACAAAGGTAGAGGCAATTCATTAATTTAGCACTTTTCAATA	720
Qy	1047	ATAGATTT 1054	
Db	721	ACACAAAT 728	

LOCUS	DEFINITION	931 bp	mRNA	linear	EST 22-VAY-2003
BX455674	BX455674 Homo sapiens				
BX455674	BX455674 Homo sapiens				
CS0DF034YL23	CS0DF034YL23 3-PRIME, mRNA sequence.				

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

EX455674
EX455674.1
EST.
GI:31022729

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 931)

1. Bases 1 and 931

COMMENT

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5629.f For
more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSDBAK043BH05NM1&cluster=5629.f)
[cgi-bin/cluster.cgi?seq=CSDBAK043BH05NM1&cluster=5629.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSDBAK043BH05NM1&cluster=5629.f). Contact :
Peng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Fairway Avenue Genoscope sequence ID : CSDBAK043BH05NM1.

FEATURES
SOURCE

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:3606"
/clone="CSDDP034Y123"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-Oligo(47) primer. Five prime end
enriched, double-strand cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

Query Match	57.3%	Score 683.2;	DB 13;	Length 931;
Best Local Similarity	94.3%	Pred. No. 4.5e-86;		
Matches 741; Conservative	0;	Mismatches 40;	Indels 5;	Gaps 3;

OY	230	CCACCAAGAGGATTCCTCCGACAGAGCTTCCTCAATGGGGGCATG--TGCTCCC	287
Db	847	CCCAAGAGGGAACTTCCTCTACCAAGCTTCTTCAAATGGGGCCCTGGGTCTCCAC	788
OY	288	CAGGGGGTCTTGCCCCCGCCAGAGAGGACAGGTCCA-GGCTGGGTGCTGTCAATCC	346
Db	787	CAGAGGGTCTTGCCCCCGCCAGACAGGACAGGTACAGAGGCTGGATTCTGTCAATCC	728
OY	347	CCCTTCATATGACCTC--AACCAGAGCTGGTGAACATGGGCTTCGCGAGTGGCATCTCG	404
Db	727	CCCTTCAGACCTTCACACGGGCAAGCTGGTAACATAGCTTTCGCGAGTGGCATCTCG	668
OY	405	GCACACATGCTGTGAGCCGGTGAACCTTCATCTGTCTCTCTTCCTGCTCATGATGCTTG	464
Db	667	GCACACATGCTGTGAGCCGGTGAACCTTCATCTGCTCTCATCTGCTCATGATGCTTG	608
OY	465	GGTTCGTGGCTCCTCCTCGTGTGGCCTTGATCACTGGGTGTCACACAGTCAAGCTTG	524
Db	607	GGTTCGTGGCTCCTCCTCGTGTGGCTTGATCACTGGGTGTCACACAGTCAAGCTTG	548
OY	525	GACCTCTGAGGCTGATAGGGATGGATTGTGTGAGAGGACTTGTCTGGCTTGGTGTGA	584
Db	547	GACCTCTGAGGCTGATAGGGATGGATTGTGTGAGAGGACTTGTCTGGCTTGGTGTGA	488
OY	585	GAGCAGGCATATTTTGAAGGGAATCTGGTGTGCTTGAAGTATGATCAAGAGGGGAACC	644
Db	487	GAGCAGGCATATTTGAGGGGATCTGGTGTGCTTGAAGTATGATCAAGAGGGGAACC	428
OY	645	AACGGTGTGTTCCTCTTGTGTAAAGGTGAGGCAAGGAGACGTTAGTCCAGCAT	704
Db	427	AACGGTGTGTTCCTCTTGTGTAAAGGTGAGGCAAGGAGATGTTAGTCCAGCAT	368
OY	705	TTTCAAAAGTGGGTGGGTCCGTGTGTTCCGAAATCTTTTAGGTGTATGGGGCGTGC	764
Db	367	TTTCAAAAGTGGGTGGGTCCGTGTGTTCCGAAATCTTTTAGGTGTATGGGGCGTGC	308
OY	765	ATTAAATGGCAAAATACAGACAGAAAGGAGATGCCCTTCCCAATTTCTCATCTCTT	824
Db	307	ATTAAATGGCAAAATACAGACAGAAAGGAGATGCCCTTCCCAATTTCTCATCTCTT	248

Best Local Similarity 98.0%; Pred. No. 3.3e-66; Matches 547; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AAACCGCTGAGACCGCGGAGTGAACGCGCGCGAGGCGCGGAGTGGCGCTGACAGCA 60
Db 9 AAACCGCTGAGACCGCGGAGTGAACGCGCGCGAGGCGCGGAGTGGCGCTGACAGCA 68
QY 61 CAGCATCTACTCAGCGTGGTCACTCTGTGAACATCACTGACGCAAGCTTCCCTCAAT 120
Db 69 CAGCATCTACTCAGCGTGGTCACTCTGTGAACATCACTGACGCAAGCTTCCCTCAAT 128
QY 121 TTCTGTGAGACCGCATCAAGGAGCCACAGCGCTGGGAGAGTGGTGGATCTTGGACAA 180
Db 129 TTCTGTGAGACCGCATCAAGGAGCCACAGCGCTGGGAGAGTGGTGGATCTTGGACAA 188
QY 181 TGGGGAATGTGACAGACGACGACCCCGAGTGAAGACCACTACCAAGCAAGAGAG 240
Db 189 TGGGGAATGTGACAGACGACGACCCCGAGTGAAGACCACTACCAAGCAAGAGAG 248
QY 241 TAGCATCTCCGACAGAGCTTTTCAATAGGGGCAATGGTCTCCCGAGGGGTCTTG 300
Db 249 TAGCATCTCCGACAGAGCTTTTCAATAGGGGCAATGGTCTCCCGAGGGGTCTTG 308
QY 301 CCCCAGCAGACAGGACAGGCTGAGGCTGAGTGGTCTCACTCCCTTCAATGACCT 360
Db 309 CCCCAGCAGACAGGACAGGCTGAGGCTGAGTGGTCTCACTCCCTTCAATGACCT 368
QY 361 CAACCGGACGCTGTGAACATGGGCTTTCCGAGTGGCATCTGCGAACCATGCTGTGA 420
Db 369 CAACCGGACGCTGTGAACATGGGCTTTCCGAGTGGCATCTGCGAACCATGCTGTGA 428
QY 421 GCGGAGTACCTCATCTGCTCTCTTCTGCTCATGATGCTGTGTTGAGGCTCTCT 480
Db 429 GCGGAGTACCTCATCTGCTCTCTTCTGCTCATGATGCTGTGTTGAGGCTCTCT 488
QY 481 CTTGTTGGCTTGTCTACTGCTGTGTCACCTGAGTCAAGCGGTGAAGCTTGAAGGCTGA 540
Db 489 CTTGTTGGCTTGTCTACTGCTGTGTCACCTGAGTCAAGCGGTGAAGCTTGAAGGCTGA 548
QY 541 TAGGGGTGGGTTGTGTA 558
Db 549 TAGGGGTGGGCTGTGTA 566

RESULT 14
CA942743/c 541 bp mRNA linear EST 30-DEC-2002
LOCUS 1166105.X1 HR85 islet Homo sapiens cDNA clone IMAGE:6607449 3'
DEFINITION mRNA sequence.

ACCESSION CA942743
VERSION CA942743.1 GI:27431223
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Butleria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 541)
Melton, D., Brown, J., Kenty, G., Pernutt, A., Lee, C., Kaestner, K.,
Lemihka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hiller, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blstein, A.,
Schmitt, A., Theising, B., Ritzer, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, K., Tsagarelis, R.,
Williams, T., Jackson, Y., and Bowers, Y.

TITLE
JOURNAL
COMMENT
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: 1166105.Y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 455.
Location/Qualifiers

FEATURES

source
1..541
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6607449"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site: 1;
Notif. Site 2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Pernutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 45.0%; Score 536.2; DB 14; Length 541;
Best Local Similarity 99.4%; Pred. No. 1.9e-65;
Matches 538; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 531 TGAAGGCGTGAATAGGGGCTGTTGTTGAGAGGAGCTTCTGAGGCTTGTGTGAGAGAG 590
Db 541 TGAAGGCGTGAATAGGGGCTGTTGTTGAGAGGAGCTTCTGAGGCTTGTGTGAGAGAG 482
QY 591 GCATATTTGAGAGGAGATCTGCTGCTTGAAGATGATCAGAGAGGAGACACAGGT 650
Db 481 GCATATTTGAGAGGAGATCTGCTGCTTGAAGATGATCAGAGAGGAGACACAGGT 422
QY 651 GTGTGTTTCCCTTGTGTTTAAAGGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 710
Db 421 GTGTGTTTCCCTTGTGTTTAAAGGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 362
QY 711 AGTGTGGGTGGGTCGTTGTTCCGAGATCTTTAGTGGTGTAGGGGCTGCAATTAG 770
Db 361 AGTGTGGGTGGGTCGTTGTTCCGAGATCTTTAGTGGTGTAGGGGCTGCAATTAG 302
QY 771 TGGCAGAAATCAGAGCAAGAAAGGAGTGCCTTCCCAATTCTCTCAATCTTTATGCG 830
Db 301 TGGCAGAAATCAGAGCAAGAAAGGAGTGCCTTCCCAATTCTCTCAATCTTTATGCG 242
QY 831 GAGAGATCTCAGCTGATGCGCAATGTCGATGCTGTGAGAGATGCGGAGCTCT 890
Db 241 GAGAGATCTCAGCTGATGCGCAATGTCGATGCTGTGAGAGATGCGGAGCTCT 182
QY 891 CCTGTGCTTGAAGGAGACGACTTGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 950
Db 181 CCTGTGCTTGAAGGAGACGACTTGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
QY 951 AGCATGTTTGTGTTTAAATGATATTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1010
Db 121 AGCATGTTTGTGTTTAAATGATATTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 62
QY 1011 GTAGTGCAATCTATTATGAGCTTTTCAATTAATAGATTTAAGTAAAAAATGAAAAA 1070
Db 61 GTAGTGCAATCTATTATGAGCTTTTCAATTAATAGATTTAAGTAAAAAATGAAAAA 2
QY 1071 A 1071
Db 1 A 1

RESULT 15

BUT733254/c		533 bp	mRNA	linear	EST 09-OCT-2002
LOCUS	BUT733254				
DEFINITION	UI-E-C11-afs-m-18-0-UI.s1 UI-E-C11 Homo sapiens CDNA clone				
ACCESSION	UI-E-C11-afs-m-18-0-UI 3'				
VERSION	BUT733254.1				
KEYWORDS	BUT733254.1 GI:23659969				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				

OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 00:20:26 ; Search time 117 Seconds

5653.859 Million cell updates/sec

Title: US-10-001-885-31
Postcard count: 1103

Sequence: 1 aaacgcctggagccgcgggaaagagacacacggaacc 1192

Scoring table: IDENTITY_NUC

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum	DB	seq	length:	0
Maximum	DB	seq	length:	3

Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*

- ```

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

| Query Length DB |       |       |        |    | ID                  | Description          |
|-----------------|-------|-------|--------|----|---------------------|----------------------|
| No.             | Score | Query | Length | DB |                     |                      |
| 1               | 1012  | 84.9  | 1049   | 4  | US-09-149-476-58    | Sequence 50, App1    |
| 2               | 473   | 39.7  | 483    | 4  | US-09-621-976-864   | Sequence 864, App1   |
| 3               | 391.8 | 32.9  | 519    | 4  | US-09-621-976-1064  | Sequence 1064, App1  |
| 4               | 222.8 | 19.5  | 364    | 4  | US-09-621-976-17202 | Sequence 17202, App1 |
| 5               | 47.2  | 4.0   | 1441   | 3  | US-08-831-994-63    | Sequence 63, App1    |
| 6               | 45.8  | 3.8   | 1971   | 2  | US-08-832-770-4     | Sequence 4, App1     |
| 7               | 45.6  | 3.8   | 664    | 4  | US-09-904-615-66    | Sequence 66, App1    |
| 8               | 45.4  | 3.8   | 2409   | 3  | US-09-233-332C-8    | Sequence 8, App1     |
| 9               | 45.4  | 3.8   | 2409   | 4  | US-09-839-497A-8    | Sequence 8, App1     |
| 10              | 45.2  | 3.8   | 2485   | 4  | US-09-889-463A-9    | Sequence 9, App1     |
| 11              | 45.2  | 3.8   | 3138   | 1  | US-07-867-106-4     | Sequence 4, App1     |
| 12              | 45.2  | 3.8   | 640681 | 4  | US-09-750-988-1     | Sequence 1, App1     |
| 13              | 45    | 3.8   | 3994   | 4  | US-09-728-946-7     | Sequence 7, App1     |
| 14              | 44.8  | 3.8   | 2381   | 1  | US-08-021-608D-9    | Sequence 9, App1     |
| 15              | 44.8  | 3.8   | 2381   | 1  | US-08-726-160-9     | Sequence 9, App1     |
| 16              | 44.8  | 3.8   | 2381   | 5  | PCT-US94-01782-9    | Sequence 9, App1     |
| 17              | 44.8  | 3.8   | 2384   | 1  | US-08-021-608D-1    | Sequence 1, App1     |
| 18              | 44.8  | 3.8   | 2384   | 1  | US-08-726-160-1     | Sequence 1, App1     |
| 19              | 44.8  | 3.8   | 2384   | 5  | PCT-US94-01782-1    | Sequence 1, App1     |
| 20              | 44.6  | 3.7   | 1466   | 4  | US-08-984-919A-10   | Sequence 10, App1    |
| 21              | 44.6  | 3.7   | 1466   | 4  | US-08-984-919A-12   | Sequence 12, App1    |
| 22              | 44.6  | 3.7   | 1472   | 3  | US-08-761-420-10    | Sequence 10, App1    |
| 23              | 44.6  | 3.7   | 1472   | 3  | US-08-761-420-12    | Sequence 12, App1    |
| 24              | 44.6  | 3.7   | 1472   | 4  | US-08-874-102-10    | Sequence 10, App1    |
| 25              | 44.6  | 3.7   | 1472   | 4  | US-08-874-102-12    | Sequence 12, App1    |
| 26              | 44.6  | 3.7   | 1472   | 4  | US-09-006-555A-10   | Sequence 10, App1    |
| 27              | 44.6  | 3.7   | 1472   | 4  | US-09-006-595A-12   | Sequence 12, App1    |

|   |    |      |     |      |   |                     |                     |
|---|----|------|-----|------|---|---------------------|---------------------|
| C | 28 | 44.6 | 3.7 | 1875 | 4 | US-08-984-9199-46   | Sequence 46, App1   |
| C | 29 | 44.6 | 3.7 | 1875 | 4 | US-08-984-9199-46   | Sequence 48, App1   |
| C | 30 | 44.6 | 3.7 | 1881 | 4 | US-08-974-102-48    | Sequence 46, App1   |
| C | 31 | 44.6 | 3.7 | 1881 | 4 | US-08-974-102-48    | Sequence 48, App1   |
| C | 32 | 44.6 | 3.7 | 6638 | 2 | US-08-070-301-2     | Sequence 2, App1    |
| C | 33 | 44.6 | 3.7 | 8319 | 4 | US-09-328-4756-2236 | Sequence 216, App1  |
| C | 34 | 44.4 | 3.7 | 1049 | 4 | US-09-800-729-67    | Sequence 67, App1   |
| C | 35 | 44.4 | 3.7 | 1313 | 4 | US-09-449-476-112   | Sequence 112, App1  |
| C | 36 | 44.2 | 3.7 | 336  | 4 | US-09-640-173-33    | Sequence 33, App1   |
| C | 37 | 44.2 | 3.7 | 366  | 4 | US-09-713-550-33    | Sequence 33, App1   |
| C | 38 | 44.2 | 3.7 | 933  | 3 | US-09-004-731-66    | Sequence 66, App1   |
| C | 39 | 44.2 | 3.7 | 933  | 3 | US-08-749-699-66    | Sequence 66, App1   |
| C | 40 | 44.2 | 3.7 | 933  | 3 | US-09-004-729-66    | Sequence 66, App1   |
| C | 41 | 44   | 3.7 | 154  | 4 | US-09-621-976-7748  | Sequence 7748, App1 |
| C | 42 | 44   | 3.7 | 154  | 4 | US-09-621-976-7748  | Sequence 7782, App1 |
| C | 43 | 44   | 3.7 | 593  | 4 | US-09-904-615-59    | Sequence 59, App1   |
| C | 44 | 44   | 3.7 | 593  | 4 | US-09-904-615-59    | Sequence 59, App1   |
| C | 45 | 43.8 | 3.7 | 923  | 3 | US-08-906-769-144   | Sequence 144, App1  |

## ALIGNMENTS

RESULT 1  
US-09-149-476-58  
Sequence 58, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P200291  
CURRENT APPLICATION NUMBER: US/09/149, 476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US99/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23



EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 84.9%; Score 1012; DB 4; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1.2e-263;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 46 TGGCGCTGACAGACAGACATCTACAGCGTGGTCACTCTGTGAACATCATCTAGCTG 105  
DB 1 TGGCGCTGACAGACAGACATCTACAGCGTGGTCACTCTGTGAACATCATCTAGCTG 60  
QY 106 CAGGCTCCCTCAATTTCTGTGTGACAGCCATCAAGGACCCACAGGCTGGAGATG 165  
DB 61 CAGGCTCCCTCAATTTCTGTGTGACAGCCATCAAGGACCCACAGGCTGGAGATG 120  
QY 166 GGGGATCTGGGCAATGGGGAAATGTGACAGACAGACCCCGAGTGAAGGACCACTAC 225  
DB 121 GGGGATCTGGGCAATGGGGAAATGTGACAGACAGACCCCGAGTGAAGGACCACTAC 180  
QY 226 CCAGCCACCAAGAGTATGACATCTCTGACAGACCTTCAATAGAGGACATGTGCTCC 285  
DB 181 CCAGCCACCAAGAGTATGACATCTCTGACAGACCTTCAATAGAGGACATGTGCTCC 240  
QY 286 CCAGAGGGGCTCTGAGGCGCCGCGACAGAGAGGAGTGCAGAGGCTGGTCTCAGTC 345  
DB 241 CCAGAGGGGCTCTGAGGCGCCGCGACAGAGAGGAGTGCAGAGGCTGGTCTCAGTC 300  
QY 346 CCGCTTCAATGACCTCAACCGGACGTGTGAACATGAGGCTTCCGACAGTGCATCTCG 405  
DB 301 CCGCTTCAATGACCTCAACCGGACGTGTGAACATGAGGCTTCCGACAGTGCATCTCG 360  
QY 406 CAACCATGTGTGAGACCGGTGACCTTCACTCTGCTCTCTTCTGCTCATGATGCTTGG 465  
DB 361 CAACCATGTGTGAGACCGGTGACCTTCACTCTGCTCTCTTCTGCTCATGATGCTTGG 420  
QY 466 TGGTGGTGGCTCTCTGCTGGTGGCTTGTCACTGAGTCCGACCTGAGTCAAGGCTG 525  
DB 421 TGGTGGTGGCTCTCTGCTGGTGGCTTGTCACTGAGTCCGACCTGAGTCAAGGCTG 480  
QY 526 ACCTCTGAGGCTGATAGGGGTGGTGTGTGAAGAGAACTTGTGGGCTTGTGTGAG 585  
DB 481 ACCTCTGAGGCTGATAGGGGTGGTGTGTGAAGAGAACTTGTGGGCTTGTGTGAG 540  
QY 586 AGCAGGCAATTTGAGAGGAGATCTGTGTGCTTGAAGTATGATCAAGAGGGAGCA 645  
DB 541 AGCAGGCAATTTGAGAGGAGATCTGTGTGCTTGAAGTATGATCAAGAGGGAGCA 600  
QY 646 CAGGTGTGTCTCCCTTGTGTGAAGCTGAGGAGAGAGAGCTTGTAGTCC--AGCA 703  
DB 601 CAGGTGTGTCTCCCTTGTGTGAAGCTGAGGAGAGAGAGCTTGTAGTCC--AGCA 660  
QY 704 TTTCCAAAGTGTGGTGGTCCGTTGGTCCGAGATCTTTTAAAGTGTATGGGCTG 763  
DB 661 TTTCCAAAGTGTGGTGGTCCGTTGGTCCGAGATCTTTTAAAGTGTATGGGCTG 720  
QY 764 CATTAAGTGGCAAAATCAGAGCAAGAAAGAGATGCCCTTCCCAATTTCTTCAATCTT 823  
DB 721 CATTAAGTGGCAAAATCAGAGCAAGAAAGAGATGCCCTTCCCAATTTCTTCAATCTT 780  
QY 824 TTAATGCGAAGATCTCAGCTGATGCCAAATGTTCCGATCCCTGTGAAAGACATGCC 883  
DB 781 TTAATGCGAAGATCTCAGCTGATGCCAAATGTTCCGATCCCTGTGAAAGACATGCC 840  
QY 884 GAGGTCTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 943  
DB 841 GAGGTCTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
QY 944 CTTTATGACATGTTTGTATTAATGATATTTTATTTGCTACTTATTTATTTAGG 1003

DB 901 CTTTATGACATGTTTGTATTAATGATATTTTATTTGCTACTTATTTAGG 960  
QY 1004 ACAAGGTATGAGGATCTTATTTATTTATGACCTTTTCAATTAATAGATTTAGTAAAA 1063  
DB 961 ACAAGGTATGAGGATCTTATTTATTTATGACCTTTTCAATTAATAGATTTAGTAAAA 1020  
QY 1064 AAAAAAAAAAAAAA 1077  
DB 1021 AAAAAAAAAAAAAA 1034

RESULT 2  
US-09-621-976-864

Sequence 864, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Mline Edwards, J.B.  
APPLICANT: Joberet, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT FILING DATE: 2000-07-21  
CURRENT APPLICATION NUMBER: US/09/621,976  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 864  
LENGTH: 483  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURES:  
NAME/KEY: CDS  
LOCATION: 171..482  
US-09-621-976-864

Query Match 39.7%; Score 473; DB 4; Length 483;  
Best Local Similarity 100.0%; Pred. No. 3.5e-118;

Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACCGCTGAGCGCCGCGAGTGAACGCCCGAGGCCGAGAGTGGCGCTGACAGCA 60  
DB 11 AAACCGCTGAGCGCCGCGAGTGAACGCCCGAGGCCGAGAGTGGCGCTGACAGCA 70  
QY 61 CAGCATCTACTACAGGTGGGTCACTCTGTGAACATCACTGACAGCTTCCCTCAAT 120  
DB 71 CAGCATCTACTACAGGTGGGTCACTCTGTGAACATCACTGACAGCTTCCCTCAAT 130  
QY 121 TTTGTGTGACAGCCATCAGGAGCCCAAGAGCGCTGGAGAGATGGTGGAGATCTTGGCCAA 180  
DB 131 TTTGTGTGACAGCCATCAGGAGCCCAAGAGCGCTGGAGAGATGGTGGAGATCTTGGCCAA 190  
QY 181 TGGGAAATCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
DB 191 TGGGAAATCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250  
QY 241 TAGCATCTCTGACAGAGCTTCTCAATAGGGGCAATGTGCTCCCGAGGGGCTCTGG 300  
DB 251 TAGCATCTCTGACAGAGCTTCTCAATAGGGGCAATGTGCTCCCGAGGGGCTCTGG 310  
QY 301 CCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
DB 311 CCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370  
QY 361 CAACCGGAGAGGTGTGAACATGGGCTTCCGCAAGTGGATCTCGGCAACATGTGTGGA 420  
DB 371 CAACCGGAGAGGTGTGAACATGGGCTTCCGCAAGTGGATCTCGGCAACATGTGTGGA 430  
QY 421 GCCGGTGAACCTTCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 473  
DB 431 GCCGGTGAACCTTCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 483

RESULT 3  
US-09-621-976-3064

Sequence 3064, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 3064  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 93..431  
US-09-621-976-3064

Query Match 32.9%; Score 391.8; DB 4; Length 519;  
Best Local Similarity 98.8%; Pred. No. 3.2e-96;  
Matches 404; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 54 GCAGACACAGCATCTACTAGCGTGGTCACTCTGTGAACATCACTGACAGCTTC 113  
DB 112 GAATACACACACTACTACTAGCGTGGTCACTCTGTGAACATCACTGACAGCTTC 171  
QY 114 CTTCAATTTTGTGTGAGCCCATCAGGATCCCAACGCGCTGGAGAGTGGCGGATCT 173  
DB 172 CTTCAATTTTGTGTGAGCCCATCAGGATCCCAACGCGCTGGAGAGTGGCGGATCT 231  
QY 174 TGGCCATGAGGAAATGTCAGACAGACAGACCCCGAGTGAAGACCACTACCCAGCCAC 233  
DB 232 TGGCCATGAGGAAATGTCAGACAGACAGACCCCGAGTGAAGACCACTACCCAGCCAC 291  
QY 234 CAAGAGTACATCTCTCGACAGAGCTTTCTCAATGAGGCGCATGTGCTCCCGAGGG 293  
DB 292 CAAGAGTACATCTCTCGACAGAGCTTTCTCAATGAGGCGCATGTGCTCCCGAGGG 350  
QY 294 GTCCGTGAGCCCGCAGACAGACAGAGTGGTGGTGTCTCAAGTCCCGCTTCA 353  
DB 351 GTCCGTGAGCCCGCAGACAGAGTGGTGGTGTCTCAAGTCCCGCTTCA 410  
QY 354 ATGACTCAACCGGAGAGCTGTGAACATGAGGCTTTCCGAGTGGCATCTCGGCAACATG 413  
DB 411 ATGACTCAACCGGAGAGCTGTGAACATGAGGCTTTCCGAGTGGCATCTCGGCAACATG 470  
QY 414 CTGTGAGCGGTGAGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 462  
DB 471 CTGTGAGCGGTGAGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 519

RESULT 4  
US-09-621-976-17202  
Sequence 17202, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 17202  
LENGTH: 364  
TYPE: DNA  
ORGANISM: Homo sapiens

Query Match 19.5%; Score 232.8; DB 4; Length 364;  
Best Local Similarity 36.5%; Pred. No. 2.4e-53;  
Matches 122; Conservative 207; Mismatches 4; Indels 1; Gaps 1;

QY 1 AAACCGCTGAGCGCCCGGAGTGAACGCGCCGAGAGCCCGAGTGCCTCTGACACA 60  
DB 1 AARYSGSKGAPRCGCCCKGAGYGMCKSRYSRBSRSCGSGMGSCKSRMSRCR 60  
QY 61 CAGCATCTACCAACGCGGTGCTGTGAACATCACTGACGACGCAAGCTCCCTCAAT 120  
DB 61 CMKSWMSMMYMRSMKTRSTCACTKGGACMTCTGAMRYASTGMCYSTIMARY 120  
QY 121 TTCTGTGACAGCCCATCAGGAGCCACAGCGCTGTGAGAGATGTGCGATCTTGCCAA 180  
DB 121 YTCYSKXRMKCYKXKRSRRCMCMCAGSGMCSRSRAGSRYSKSGRGMWYMKGCSRA 180  
QY 181 TGGGAAATCGTGAAGACAGACAGACCCCGAGTGAAGACCACTACCAACCAAGAG 240  
DB 181 TSKGRMMWMMKGGRRRATSRYMMSMYGASKMSMCSASTMSASCMWMTMMSAGS 240  
QY 241 TAGCATCTCTGACAGAGCTTCTCAATAGGAGGCGCATGTGCTCCCGAGAGG-TCCTG 259  
DB 241 YASCAWCKSKTRRCAKMSCTYSWTRBMSKSKYCAWSRRGSKCYYSRSGSKCYCWS 300  
QY 300 GCCCGCCGACAGACAGGAGGTGCCAGGCTGG 333  
DB 301 GSCCGCCGACAGACAGGAGGTGCCAGGCTGG 334

RESULT 5  
US-08-821-994-63  
Sequence 63, Application US/08821994A  
Patent No. 6228643  
GENERAL INFORMATION:  
APPLICANT: Greenland, Andrew J  
APPLICANT: Thomas, Didier RP  
APPLICANT: Jepson, Ian  
TITLE OF INVENTION: Promoters  
FILE REFERENCE: PPD 50108  
CURRENT APPLICATION NUMBER: US/08/821,994A  
CURRENT FILING DATE: 1997-03-22  
EARLIER APPLICATION NUMBER: PCT/GB97/00729  
EARLIER FILING DATE: 1997-03-18  
EARLIER APPLICATION NUMBER: GB 9606062.9  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 63  
LENGTH: 1441  
TYPE: DNA  
ORGANISM: Brassica napus  
US-08-821-994-63

Query Match 4.0%; Score 47.2; DB 3; Length 1441;  
Best Local Similarity 58.6%; Pred. No. 0.0076;  
Matches 82; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 946 TTTTACGACGTTTGTGTTATATGATATATTTTATGCTACTTATTTATGTTAGAC 1005  
DB 1289 TGTGTCAATATATTTCAATCAAGAAATTTGCTTGTGATGATATTAAG 1348  
QY 1006 AAGTGTAGTGGCATCTATTTATTTGACCTTTCAATTAATAGATTATTAAGTAAAAA 1065  
DB 1349 AAATATATTAATAATATATTTTCTCAAAAAAATATATATATATATATATATATATAT 1408  
QY 1066 AAAAAAAAAAAAAAAAAA 1085  
DB 1409 AAAAAAAAAAAAAAAAAA 1428

Sequence 4, Application US/08892770  
Patent No. 5962670  
GENERAL INFORMATION:  
APPLICANT: Walling, Linda L.  
APPLICANT: Pautot, Veronique  
APPLICANT: Gu, Yong-Qiang  
APPLICANT: Chao, Wan Shaw  
TITLE OF INVENTION: Improved Promoters for Enhancing Plant  
TITLE OF INVENTION: Productivity  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/892,770  
FILING DATE: 15-JUL-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Quine, Jonathan A.  
REGISTRATION NUMBER: P-41,261  
REFERENCE/DOCKET NUMBER: 023070-072100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1971 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..1971  
OTHER INFORMATION: /note="tomato acidic leucine  
OTHER INFORMATION: aminopeptidase 2 (Lap2) cDNA"  
US-08-892-770-4  
Query Match 3.8%; Score 45.8; DB 2; Length 1971;  
Best Local Similarity 55.3%; Pred. No. 0.022; Indels 0; Gaps 0;  
Matches 89; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 925 AGGTGAAAAAATTCAGACTTTTGAACCTGTTTGTATGATATTTTAT 984  
DB 1806 AGATGAATAAATGCTTAATTAATTTGCACTTACCTTTCAATGAATTTATTT 1865  
QY 985 GGCTACTTATTTGTTAGACCAAGTGAGTGCATTTCTATTTATGTGACCTTTGAT 1044  
DB 1866 CCGCTCTATTATTCGATGCTGATGAGATTAATTCCTCTAGATGTAACCTTGAG 1925  
QY 1045 AATAGATTATTAAGTAAAAAATTTTAAAAAATTTTAAAAA 1085  
DB 1926 ATGAAGTGTATAAAAAATTTTAAAAAATTTTAAAAA 1966  
RESULT 7  
US-09-904-615-66  
Sequence 66, Application US/09904615  
Patent No. 656325  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 49 Human Secreted Proteins  
FILE REFERENCE: P2032PI  
CURRENT APPLICATION NUMBER: US/09/904,615

CURRENT FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 09/511,554  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 60/097,917  
PRIOR FILING DATE: 1998-08-25  
PRIOR APPLICATION NUMBER: 60/098,634  
PRIOR FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 170  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 66  
LENGTH: 664  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (31)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (63)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-904-615-66  
Query Match 3.8%; Score 45.6; DB 4; Length 664;  
Best Local Similarity 55.8%; Pred. No. 0.013; Indels 0; Gaps 0;  
Matches 87; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 930 GAAAAAATTCAGACTTTTGAACCTGTTTGTATGATATTTTATGAGCTA 989  
DB 493 GAAAAAATTTATATATGATGAGTTTGTCTTTTATTTTATTTTGGTATA 552  
QY 990 CTTATGTTTGAACCAAGTGAGTGCATTTCTATTTATGTGACCTTTCAATAATA 1049  
DB 553 TTGCTGATCTACTTAATCTTCAGAAATTAACGTTATATGAAAAA 612  
QY 1050 GATTAAAGTAAAAAATTTTAAAAAATTTTAAAAA 1085  
DB 613 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 648  
RESULT 8  
US-09-293-322C-8  
Sequence 8, Application US/09293322C  
Patent No. 6232110  
GENERAL INFORMATION:  
APPLICANT: Pallas, David C  
APPLICANT: Du, Xianxing  
TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase,  
Patent No. 6232110  
TITLE OF INVENTION: Recombinant DNA Molecules and Methods  
FILE REFERENCE: 105-97  
CURRENT APPLICATION NUMBER: US/09/293,322C  
CURRENT FILING DATE: 1999-04-16  
PRIOR APPLICATION NUMBER: US 60/082,202  
PRIOR FILING DATE: 1998-04-17  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 8  
LENGTH: 2409  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(2409)  
OTHER INFORMATION: N is A, T, G or C.  
US-09-293-322C-8  
Query Match 3.8%; Score 45.4; DB 3; Length 2409;  
Best Local Similarity 65.0%; Pred. No. 0.031; Indels 0; Gaps 0;  
Matches 67; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
QY 983 TTGGCTACTTATTTGTTAGACCAAGTGAGTGCATTTCTATTTATGTGACCTTTTCA 1042  
DB 2294 TTGTTCTTTGTTATTTATGATCTGTGTTAAAGAAAAATATATCTCCACCTTTAAA 2353

QY 1043 ATAAATGATTAAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1085  
DB 2354 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2396

## RESULT 9

US-09-839-497A-8  
Sequence 8, Application US/09839497A  
Patent No. 6528295  
GENERAL INFORMATION:  
APPLICANT: Pallas, David C.  
APPLICANT: Du, Xianxing  
TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase,  
Patent No. 6528295  
TITLE OF INVENTION: Recombinant DNA Molecules and Methods  
FILE REFERENCES: Docket No. 6528295, 105-97A  
CURRENT APPLICATION NUMBER: US/09/839,497A  
CURRENT FILING DATE: 2001-04-20  
PRIOR APPLICATION NUMBER: 60/082,202  
PRIOR FILING DATE: 1998-04-17  
PRIOR APPLICATION NUMBER: 09/293,322  
PRIOR FILING DATE: 1999-04-16  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 8  
LENGTH: 2409  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(2409)  
OTHER INFORMATION: N is A, T, G or C.  
US-09-839-497A-8

Query Match 3.8%; Score 45.4; DB 4; Length 2409;  
Best Local Similarity 65.0%; Pred. No. 0.031;  
Matches 67; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 983 TTGGCTACTTATTTAGTACAGAGTGTAGTGGCATTTCTATTATTGTGACCTTTTCA 1042  
DB 2294 TTGGTTCTTTGTTATTATGATCTGTTTAAAGAAATATATCTCCCAACCTTTAA 2353  
QY 1043 ATAAATGATTAAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1085  
DB 2354 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2396

## RESULT 10

US-09-889-463A-9  
Sequence 9, Application US/09889463A  
Patent No. 6680185  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Falco, Savetio C.  
APPLICANT: Kinney, Anthony J.  
APPLICANT: Mao, Guo-Hua  
TITLE OF INVENTION: Plant Polyphenol Oxidase Homologs  
FILE REFERENCES: B81330  
CURRENT APPLICATION NUMBER: US/09/889,463A  
CURRENT FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/119,590  
PRIOR FILING DATE: 1999-02-10  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 9  
LENGTH: 2485  
TYPE: DNA  
ORGANISM: Glycine max  
US-09-889-463A-9

Query Match 3.8%; Score 45.2; DB 4; Length 2485;  
Best Local Similarity 55.8%; Pred. No. 0.036;

Matches 86; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
QY 932 AAAATTCGACCTTTTGGACTGTTTGTGTTAAGTATTTTATGGCTACT 991  
DB 1947 ATATPAAGTAGTCTTGATGATTTGTTGCTTTATATGAAAAATAGTTTGTAT 2006  
QY 992 TTATGTTAGACAGTGTAGTGCATTTCTATTATGACCTTTCAATPAATGA 1051  
DB 2007 GTATTTTCATGACATCATPAAGATGTCCTATATGAGTGAAGAAAAA 2066  
QY 1052 TTTAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1085  
DB 2067 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2100

## RESULT 11

US-07-867-106-4  
Sequence 4, Application US/07867106  
Patent No. 5389526  
GENERAL INFORMATION:  
APPLICANT: Slade, Martin B  
APPLICANT: Chang, Andy C M  
APPLICANT: Williams, Keith L  
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular  
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris  
STREET: One Liberty Place 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/867,106  
FILING DATE: 19920625  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PJ 7187  
APPLICATION NUMBER: PCT/AU90/00530  
FILING DATE: 02-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Feeney, Joanne Longo  
REGISTRATION NUMBER: 35,134  
REFERENCE/DOCKET NUMBER: RICE-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3439  
TELEFAX: 215-568-3100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3138 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
ANTI-SENSE: NO  
US-07-867-106-4

Query Match 3.8%; Score 45.2; DB 1; Length 3138;  
Best Local Similarity 54.9%; Pred. No. 0.041;  
Matches 89; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 924 CAGGTGAAAAAATTCAGACTTTTGTAGACCTGTTTGTGTTAAGTAAATTTTAT 983  
DB 1912 CAAATTAATAAAAAAAAAATGGAATCAAAAATTTTTTTTTTTTTTTTTTTT 1971  
QY 984 TTGGCTACTTATTTTGAAGCAAGTGTAGTGCATTTCTATTATGACCTTTTCA 1043  
DB 1972 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2031

```

? LENGTH: 2381
? TYPE: Nucleic acid
? STRANDEDNESS: Double
? TOPOLOGY: Unknown
? MOLECULE TYPE: cDNA
? HYPOTHETICAL: NO
? ANTI-SENSE: YES
? ORIGINAL SOURCE:
? ORGANISM: Human
? CELL LINE: HL60
? FEATURE:

```





```

RESULT 1
US-10-001-885-31
; Sequence 31, Application US/10001885
; Publication No. US20040058319A1
GENERAL INFORMATION:
APPLICANT: Salceda, Susana
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Herive
APPLICANT: Caffferkey, Robert
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
FILE REFERENCE: DEX-0279
CURRENT APPLICATION NUMBER: US/10/001,885
PRIORITY FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/252,061
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: 60/253,257
PRIOR FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 167
SOFTWARE: PatentIn version 3.1
SEQ ID NO 31
LENGTH: 1192
TYPE: DNA
ORGANISM: Homo sapien
US-10-001-885-31

Query Match 100.0%; Score 1192; DB 13; Length 1192;
Best Local Similarity 100.0%; Pred. No. 6,9e-309;
Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACCGCCTGACCCGCGCGGAGTGAACGCGCGCGAGAGTCCGCCCTGACAGACA 60
Db 1 AAACCGCCTGACCCGCGCGGAGTGAACGCGCGCGAGAGTCCGCCCTGACAGACA 60
QY 61 CAGCATCTATCGAGCGGTGGGTGACCTCTGTGAACATCAGCATGGAAGCTCCCTCAAT 120

```

Db 61 CAGCATCTACTCAGCGTGGTCACTCTGTGAAATCACTAGTCGACGCTCCCTCAAT 120  
Qy 121 TTCTGGTGCAGCCCATCAGGGAACCAAGAGCTTGGAGGATGTGTGCGATCTTGGCCAA 180  
Db 121 TTCTGGTGCAGCCCATCAGGGAACCAAGAGCTTGGAGGATGTGTGCGATCTTGGCCAA 180  
Qy 181 TGGGAAATGTGTCAGAGCAGACGACCCCGAGTGAACCACTACCCAGCCCAAGAG 240  
Db 181 TGGGAAATGTGTCAGAGCAGACGACCCCGAGTGAACCACTACCCAGCCCAAGAG 240  
Qy 241 TACCATTTCCCTGACAGAGCTTCTTCAATAGGGGCGATGTGGCTCCCGAGGGGGCTGG 300  
Db 241 TACCATTTCCCTGACAGAGCTTCTTCAATAGGGGCGATGTGGCTCCCGAGGGGGCTGG 300  
Qy 301 CCCCCGACAGCAGGAGAGTGCAGGCTGGTGTCTAGTCCCTTCAATGACCT 360  
Db 301 CCCCCGACAGCAGGAGAGTGCAGGCTGGTGTCTAGTCCCTTCAATGACCT 360  
Qy 361 CAACCCGACAGTGTGAAATGGGCTTTCCGAGTGGCATCTTGGGACCAATGCTGTGA 420  
Db 361 CAACCCGACAGTGTGAAATGGGCTTTCCGAGTGGCATCTTGGGACCAATGCTGTGA 420  
Qy 421 GCGGTGACCTGCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
Db 421 GCGGTGACCTGCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
Qy 481 CTTGTTGGCTTGTCTACTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
Db 481 CTTGTTGGCTTGTCTACTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
Qy 541 TAAAGGGGGTGTGTGTGAGAGGAGCTTGGTGGGCTTGTGTGAGAGGAGCATATTGG 600  
Db 541 TAAAGGGGGTGTGTGTGAGAGGAGCTTGGTGGGCTTGTGTGAGAGGAGCATATTGG 600  
Qy 601 AGGGGATCTGGTGTGTGTGAGAGGAGCTTGTGTGAGAGGAGCATATTGG 660  
Db 601 AGGGGATCTGGTGTGTGTGAGAGGAGCTTGTGTGAGAGGAGCATATTGG 660  
Qy 661 CTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720  
Db 661 CTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720  
Qy 721 GGTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
Db 721 GGTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
Qy 781 TCAAGAGCAAGAACGATGCTTCCCAATCTCTCAATCTCTTATGCGGAGAGATCT 840  
Db 781 TCAAGAGCAAGAACGATGCTTCCCAATCTCTCAATCTCTTATGCGGAGAGATCT 840  
Qy 841 CAGCTGATCCCAATGCTTCCCAATCTCTCAATCTCTTATGCGGAGAGATCT 900  
Db 841 CAGCTGATCCCAATGCTTCCCAATCTCTCAATCTCTTATGCGGAGAGATCT 900  
Qy 901 GGGAGCAGACTTGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
Db 901 GGGAGCAGACTTGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
Qy 961 TTTGTTTAAATGTGATTTTAAATGTGATTTTAAATGTGATTTTAAATGTGATTTT 1020  
Db 961 TTTGTTTAAATGTGATTTTAAATGTGATTTTAAATGTGATTTTAAATGTGATTTT 1020  
Qy 1021 TCTATTATGTGATTTTAAATGTGATTTTAAATGTGATTTTAAATGTGATTTTAAAT 1080  
Db 1021 TCTATTATGTGATTTTAAATGTGATTTTAAATGTGATTTTAAATGTGATTTTAAAT 1080  
Qy 1081 AAAAAGTGTGGGGTGTGATTTTAAATGTGATTTTAAATGTGATTTTAAATGTGATTT 1140  
Db 1081 AAAAAGTGTGGGGTGTGATTTTAAATGTGATTTTAAATGTGATTTTAAATGTGATTT 1140  
Qy 1141 GGGCCCAATTTCCCAATTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1192  
Db 1141 GGGCCCAATTTCCCAATTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1192

RESULT 2  
US-10-264-237-619  
; Sequence 619, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birex et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P131P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; PRIOR FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 619  
; LENGTH: 1108  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1101)..(1101)  
; OTHER INFORMATION: n equals a,t,c,g, or c  
US-10-264-237-619  
Query Match 89.0%; Score 1061.4; DB 16; Length 1108;  
Best Local Similarity 99.6%; Pred. No. 7.2e-274;  
Matches 1059; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
Qy 15 GCGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 74  
Db 1 GCGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60  
Qy 75 CGTGGTCACTCTGTGAAATCACTGACCTGACCTGACCTGACCTGACCTGACCTGAC 134  
Db 61 CGTGGTCACTCTGTGAAATCACTGACCTGACCTGACCTGACCTGACCTGACCTGAC 120  
Qy 135 ATCAGGAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 194  
Db 121 ATCAGGAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
Qy 195 AGAGCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 254  
Db 181 AGAGAGAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
Qy 255 AGAGCTTCTCAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 314  
Db 241 AGAGCTTCTCAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
Qy 315 AGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 374  
Db 301 AGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
Qy 375 TGAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 434  
Db 361 TGAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
Qy 435 TCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 494  
Db 421 TCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
Qy 495 TCTACCTGTGTCTCCCACTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 554  
Db 481 TCTACCTGTGTCTCCCACTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
Qy 555 TTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 614  
Db 541 TTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
Qy 615 TGCCTTGAAGGATGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 674



FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/809,391  
CURRENT FILING DATE: 2001-03-16  
Prior application data removed - consult PAM or file wrapper  
NUMBER OF SEQ ID NOS: 761  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 58  
LENGTH: 1049  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-809-391-58

Query Match 84.9%; Score 1012; DB 10; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1.3e-260;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 46 TCGGCGCTGAGAGACAGCATCTACAGGGGTGGTCACTTGTGAACATCACTGACTG 105  
DB 1 TCGGCGCTGAGAGACAGCATCTACAGGGGTGGTCACTTGTGAACATCACTGACTG 60  
QY 106 CAAGCCTCCCTCATTTCTGGTGAAGCCCATCAGGACCCACAGCCCTGGAGAGATGGT 165  
DB 61 CAAGCCTCCCTCATTTCTGGTGAAGCCCATCAGGACCCACAGCCCTGGAGAGATGGT 120  
QY 166 GCGGATCTGGCCCATGGGAAATCTGTGAGAGACAGACCCCGAGTGAAGCACTAC 225  
DB 121 GCGGATCTGGCCCATGGGAAATCTGTGAGAGACAGACCCCGAGTGAAGCACTAC 180  
QY 226 CCAGCCACAGAGGTGAGCATCTCTGAGAGAGCTTCAATAGGGGAGCATGATCTCC 285  
DB 181 CCAGCCACAGAGGTGAGCATCTCTGAGAGAGCTTCAATAGGGGAGCATGATCTCC 240  
QY 286 CCCAGGGGGTCTGTCGCCCGCCAGCAGCAGGAGGAGGCTGGTGGTCTGATGTC 345  
DB 241 CCCAGGGGGTCTGTCGCCCGCCAGCAGCAGGAGGAGGCTGGTGGTCTGATGTC 300  
QY 346 CCCCTTAATGACCTCAACCGGAGCTGTGAACATGAGGCTTCCGAGTGGATCTCGG 405  
DB 301 CCCCTTAATGACCTCAACCGGAGCTGTGAACATGAGGCTTCCGAGTGGATCTCGG 360  
QY 406 CAACCATGCTGTGAGAGCGGTGACCTTCATCTGCTCTTCTGCTCAATGATGCTGG 465  
DB 361 CAACCATGCTGTGAGAGCGGTGACCTTCATCTGCTCTTCTGCTCAATGATGCTGG 420  
QY 466 TGTTCGTGGGCTCTCTCTGCTGTGACCTGTGACCTGTGACCTGTGACCTGTGAC 525  
DB 421 TGTTCGTGGGCTCTCTCTGCTGTGACCTGTGACCTGTGACCTGTGACCTGTGAC 480  
QY 526 ACCTGTAGAGGCTGATGAGGGGTGGTGTGTTGAGAGGGAATTGCTGGGCTTGTGAG 585  
DB 481 ACCTGTAGAGGCTGATGAGGGGTGGTGTGTTGAGAGGGAATTGCTGGGCTTGTGAG 540  
QY 586 AGCAGGCAATTTTGAAGGGGATCTGGTGTGCTTTGAAGGATGATCAAGAGGGGACA 645  
DB 541 AGCAGGCAATTTTGAAGGGGATCTGGTGTGCTTTGAAGGATGATCAAGAGGGGACA 600  
QY 646 CAGGTGTGTTTCCCTTGTGTTAAGGTGAGGAGAGGGGAGAGCTTAACTCC--AGCA 703  
DB 601 CAGGTGTGTTTCCCTTGTGTTAAGGTGAGGAGAGGGGAGAGCTTAACTCCAGCAT 660  
QY 704 TTTCCAAAGTGTGGGTGGGTCCGTTGTTCCGAGATCTTTTGTAGTGTGGGCTG 763  
DB 661 TTTCCAAAGTGTGGGTGGGTCCGTTGTTCCGAGATCTTTTGTAGTGTGGGCTG 720  
QY 764 CATTAAGTGCACAAATGAGAGAGAAAGCATGCTTCCCAATTTCTCATCTCT 823  
DB 721 CATTAAGTGCACAAATGAGAGAGAAAGCATGCTTCCCAATTTCTCATCTCT 780  
QY 824 TTATGCGAGAGATCTCAGCTGTGATGCCAACAATGTTCCGATGCTGTGAGAGCATGCC 883  
DB 781 TTATGCGAGAGATCTCAGCTGTGATGCCAACAATGTTCCGATGCTGTGAGAGCATGCC 840

DB 841 GACGTCCTCTGCTGAGAGAGAGAGACTTGGGCTTAGGGCAGGAGGAAAAATTCAGCA 900  
QY 944 CTTTGTGACACTGTTTGTGTTTAATGATATTTTATGCTACTTATTTGTTAGG 1003  
DB 901 CTTTGTGACACTGTTTGTGTTTAATGATATTTTATGCTACTTATTTGTTAGG 960  
QY 1004 ACAAGTGTAGTGCATCTATTATTGTGACCTTTCAATTAATGATTTAAGTAAAAA 1063  
DB 961 ACAAGTGTAGTGCATCTATTATTGTGACCTTTCAATTAATGATTTAAGTAAAAA 1020

QY 1064 AAAAAAAAAAAAAA 1077  
DB 1021 AAAAAAAAAAAAAA 1034

RESULT 5  
US-09-882-171-58  
Sequence 58, Application US/09882171  
Publication No. US20030175858A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/882,171

PRIOR FILING DATE: 2001-06-18  
CURRENT FILING DATE: 2001-03-16

PRIOR APPLICATION NUMBER: 09/809,391  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 09/149,476  
PRIOR FILING DATE: 1998-09-08

PRIOR APPLICATION NUMBER: PCT/US98/04493  
PRIOR FILING DATE: 1998-03-06

PRIOR APPLICATION NUMBER: 60/040,162  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/040,333  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/038,621  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/040,626  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/040,334  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/040,336  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/040,163  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/047,600  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,615  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,597  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,502  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,633  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,583  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,617  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,618  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,503  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,592  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,581  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,584  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,500  
PRIOR FILING DATE: 1997-05-23

1 PRIOR FILING DATE: 1997-05-23  
2 PRIOR APPLICATION NUMBER: 60/047,492  
3 PRIOR FILING DATE: 1997-05-23  
4 PRIOR APPLICATION NUMBER: 60/047,598  
5 PRIOR FILING DATE: 1997-05-23  
6 PRIOR APPLICATION NUMBER: 60/047,613  
7 PRIOR FILING DATE: 1997-05-23  
8 PRIOR APPLICATION NUMBER: 60/047,582  
9 PRIOR FILING DATE: 1997-05-23  
10 PRIOR APPLICATION NUMBER: 60/047,596  
11 PRIOR FILING DATE: 1997-05-23  
12 PRIOR APPLICATION NUMBER: 60/047,612  
13 PRIOR FILING DATE: 1997-05-23  
14 PRIOR APPLICATION NUMBER: 60/047,632  
15 PRIOR FILING DATE: 1997-05-23  
16 PRIOR APPLICATION NUMBER: 60/047,601  
17 PRIOR FILING DATE: 1997-05-23  
18 PRIOR APPLICATION NUMBER: 60/043,580  
19 PRIOR FILING DATE: 1997-04-11  
20 PRIOR APPLICATION NUMBER: 60/043,568  
21 PRIOR FILING DATE: 1997-04-11  
22 PRIOR APPLICATION NUMBER: 60/043,314  
23 PRIOR FILING DATE: 1997-04-11  
24 PRIOR APPLICATION NUMBER: 60/043,569  
25 PRIOR FILING DATE: 1997-04-11  
26 PRIOR APPLICATION NUMBER: 60/043,311  
27 PRIOR FILING DATE: 1997-04-11  
28 PRIOR APPLICATION NUMBER: 60/043,671  
29 PRIOR FILING DATE: 1997-04-11  
30 PRIOR APPLICATION NUMBER: 60/043,674  
31 PRIOR FILING DATE: 1997-04-11  
32 PRIOR APPLICATION NUMBER: 60/043,669  
33 PRIOR FILING DATE: 1997-04-11  
34 PRIOR APPLICATION NUMBER: 60/043,312  
35 PRIOR FILING DATE: 1997-04-11  
36 PRIOR APPLICATION NUMBER: 60/043,313  
37 PRIOR FILING DATE: 1997-04-11  
38 PRIOR APPLICATION NUMBER: 60/043,672  
39 PRIOR FILING DATE: 1997-04-11  
40 PRIOR APPLICATION NUMBER: 60/043,315  
41 PRIOR FILING DATE: 1997-04-11  
42 PRIOR APPLICATION NUMBER: 60/048,974  
43 PRIOR FILING DATE: 1997-06-06  
44 PRIOR APPLICATION NUMBER: 60/056,886  
45 PRIOR FILING DATE: 1997-08-22  
46 PRIOR APPLICATION NUMBER: 60/056,877  
47 PRIOR FILING DATE: 1997-08-22  
48 PRIOR APPLICATION NUMBER: 60/056,889  
49 PRIOR FILING DATE: 1997-08-22  
50 PRIOR APPLICATION NUMBER: 60/056,893  
51 PRIOR FILING DATE: 1997-08-22  
52 PRIOR APPLICATION NUMBER: 60/056,630  
53 PRIOR FILING DATE: 1997-08-22  
54 PRIOR APPLICATION NUMBER: 60/056,878  
55 PRIOR FILING DATE: 1997-08-22  
56 PRIOR APPLICATION NUMBER: 60/056,662  
57 PRIOR FILING DATE: 1997-08-22  
58 PRIOR APPLICATION NUMBER: 60/056,872  
59 PRIOR FILING DATE: 1997-08-22  
60 PRIOR APPLICATION NUMBER: 60/056,882  
61 PRIOR FILING DATE: 1997-08-22  
62 PRIOR APPLICATION NUMBER: 60/056,637  
63 PRIOR FILING DATE: 1997-08-22  
64 PRIOR APPLICATION NUMBER: 60/056,903  
65 PRIOR FILING DATE: 1997-08-22  
66 PRIOR APPLICATION NUMBER: 60/056,888  
67 PRIOR FILING DATE: 1997-08-22  
68 PRIOR APPLICATION NUMBER: 60/056,879  
69 PRIOR FILING DATE: 1997-08-22  
70 PRIOR APPLICATION NUMBER: 60/056,880  
71 PRIOR FILING DATE: 1997-08-22  
72 PRIOR APPLICATION NUMBER: 60/056,894  
73 PRIOR FILING DATE: 1997-08-22

74 PRIOR APPLICATION NUMBER: 60/056,911  
75 PRIOR FILING DATE: 1997-08-22  
76 PRIOR APPLICATION NUMBER: 60/056,636  
77 PRIOR FILING DATE: 1997-08-22  
78 PRIOR APPLICATION NUMBER: 60/056,874  
79 PRIOR FILING DATE: 1997-08-22  
80 PRIOR APPLICATION NUMBER: 60/056,910  
81 PRIOR FILING DATE: 1997-08-22  
82 PRIOR APPLICATION NUMBER: 60/056,864  
83 PRIOR FILING DATE: 1997-08-22  
84 PRIOR APPLICATION NUMBER: 60/056,631  
85 PRIOR FILING DATE: 1997-08-22  
86 PRIOR APPLICATION NUMBER: 60/056,845  
87 PRIOR FILING DATE: 1997-08-22  
88 PRIOR APPLICATION NUMBER: 60/056,892  
89 PRIOR FILING DATE: 1997-08-22  
90 PRIOR APPLICATION NUMBER: 60/057,761  
91 PRIOR FILING DATE: 1997-08-22  
92 PRIOR APPLICATION NUMBER: 60/047,595  
93 PRIOR FILING DATE: 1997-05-23  
94 PRIOR APPLICATION NUMBER: 60/047,586  
95 PRIOR FILING DATE: 1997-05-23  
96 PRIOR APPLICATION NUMBER: 60/047,590  
97 PRIOR FILING DATE: 1997-05-23  
98 PRIOR APPLICATION NUMBER: 60/047,594  
99 PRIOR FILING DATE: 1997-05-23  
100 PRIOR APPLICATION NUMBER: 60/047,589  
101 PRIOR FILING DATE: 1997-05-23  
102 PRIOR APPLICATION NUMBER: 60/047,593  
103 PRIOR FILING DATE: 1997-05-23  
104 PRIOR APPLICATION NUMBER: 60/047,614  
105 PRIOR FILING DATE: 1997-05-23  
106 PRIOR APPLICATION NUMBER: 60/043,578  
107 PRIOR FILING DATE: 1997-04-11  
108 PRIOR APPLICATION NUMBER: 60/043,576  
109 PRIOR FILING DATE: 1997-04-11  
110 PRIOR APPLICATION NUMBER: 60/047,501  
111 PRIOR FILING DATE: 1997-05-23  
112 PRIOR APPLICATION NUMBER: 60/043,670  
113 PRIOR FILING DATE: 1997-04-11  
114 PRIOR APPLICATION NUMBER: 60/056,632  
115 PRIOR FILING DATE: 1997-08-22  
116 PRIOR APPLICATION NUMBER: 60/056,664  
117 PRIOR FILING DATE: 1997-08-22  
118 PRIOR APPLICATION NUMBER: 60/056,876  
119 PRIOR FILING DATE: 1997-08-22  
120 PRIOR APPLICATION NUMBER: 60/056,881  
121 PRIOR FILING DATE: 1997-08-22  
122 PRIOR APPLICATION NUMBER: 60/056,909  
123 PRIOR FILING DATE: 1997-08-22  
124 PRIOR APPLICATION NUMBER: 60/056,875  
125 PRIOR FILING DATE: 1997-08-22  
126 PRIOR APPLICATION NUMBER: 60/056,862  
127 PRIOR FILING DATE: 1997-08-22  
128 PRIOR APPLICATION NUMBER: 60/056,887  
129 PRIOR FILING DATE: 1997-08-22  
130 PRIOR APPLICATION NUMBER: 60/056,908  
131 PRIOR FILING DATE: 1997-08-22  
132 PRIOR APPLICATION NUMBER: 60/048,964  
133 PRIOR FILING DATE: 1997-06-06  
134 PRIOR APPLICATION NUMBER: 60/057,650  
135 PRIOR FILING DATE: 1997-09-05  
136 PRIOR APPLICATION NUMBER: 60/056,884  
137 PRIOR FILING DATE: 1997-08-22  
138 PRIOR APPLICATION NUMBER: 60/057,669  
139 PRIOR FILING DATE: 1997-09-05

Query Match 84.9%; Score 1012; DB 10; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1,3e-260;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

```
QY 46 TCGGCGCTGACAGACAGCATCTAAGCGTGGGTCACTGTGAAACATCACTGACTG 105
Db 1 TCGGCGCTGACAGACAGCATCTAAGCGTGGGTCACTGTGAAACATCACTGACTG 60
QY 106 CAAGCCTCCCTCAATTTCTGTGAGCCCAATCAAGGGGCCAAGCGCTGGGAGATGGT 165
Db 61 CAAGCCTCCCTCAATTTCTGTGAGCCCAATCAAGGGGCCAAGCGCTGGGAGATGGT 120
QY 166 GCGGATCTGGCCAAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 225
Db 121 GCGGATCTGGCCAAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 226 CCAAGCCCAAGAGAGATGCAATCTCTGACAGAGCTTTCAATAGGGGCCATGTGCTCC 285
Db 181 CCAAGCCCAAGAGAGATGCAATCTCTGACAGAGCTTTCAATAGGGGCCATGTGCTCC 240
QY 286 CCGAGGGGGTCTGGCCCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
Db 241 CCGAGGGGGTCTGGCCCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 346 CCGCTTCAATGACCTCAACCGGAGAGCTGTGAAATAGGGCTTTCCGAGTGGCATCTCG 405
Db 301 CCGCTTCAATGACCTCAACCGGAGAGCTGTGAAATAGGGCTTTCCGAGTGGCATCTCG 360
QY 406 CAACCAATGCTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
Db 361 CAACCAATGCTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 466 TGTTCGTGAGCTCTCTGAGTGGCTTTGTCTAAGCTGTGAGAGAGAGAGAGAGAG 525
Db 421 TGTTCGTGAGCTCTCTGAGTGGCTTTGTCTAAGCTGTGAGAGAGAGAGAGAGAG 480
QY 526 ACCTCTGAG 585
Db 481 ACCTCTGAG 540
QY 586 AGCAGAGATATTTGAG 645
Db 541 AGCAGAGATATTTGAG 600
QY 646 CAGGTGTGTCTCCCTTTGTGTAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 703
Db 601 CAGGTGTGTCTCCCTTTGTGTAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 704 TTTTCAAAAGTGGAG 763
Db 661 TTTTCAAAAGTGGAG 720
QY 764 CATTAAATGAGCAAAAATCAGAGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAG 823
Db 721 CATTAAATGAGCAAAAATCAGAGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 824 TTTATGCGGAAATATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 883
Db 781 TTTATGCGGAAATATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 884 GAGCTCTCTCTGAG 943
Db 841 GAGCTCTCTCTGAG 900
QY 944 CTTTTTATGAG 1003
Db 901 CTTTTTATGAG 960
QY 1004 ACAAGTGTGTGAG 1063
Db 961 ACAAGTGTGTGAG 1020
```

Db 1021 AAAAAAAAAAAAAA 1014

```
RESULT 6
US-10-164-861-58
; Sequence 58, Application US/10164861
; Publication No. US20030225248A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/10/164, 861
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US/09/149, 476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; NUMBER OF SEQ ID NOS: 757
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 1049
; TYPER: DNA
; ORGANISM: Homo sapiens
US-10-164-861-58
```

Query Match 84.9%; Score 1012; DB 13; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1,3e-260;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

```
QY 46 TCGGCGCTGACAGACAGCATCTAAGCGTGGGTCACTGTGAAACATCACTGACTG 105
Db 1 TCGGCGCTGACAGACAGCATCTAAGCGTGGGTCACTGTGAAACATCACTGACTG 60
QY 106 CAAGCCTCCCTCAATTTCTGTGAGCCCAATCAAGGGGCCAAGCGCTGGGAGATGGT 165
Db 61 CAAGCCTCCCTCAATTTCTGTGAGCCCAATCAAGGGGCCAAGCGCTGGGAGATGGT 120
QY 166 GCGGATCTGGCCAAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 225
Db 121 GCGGATCTGGCCAAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 226 CCAAGCCCAAGAGAGATGCAATCTCTGACAGAGCTTTCAATAGGGGCCATGTGCTCC 285
Db 181 CCAAGCCCAAGAGAGATGCAATCTCTGACAGAGCTTTCAATAGGGGCCATGTGCTCC 240
QY 286 CCGAGGGGGTCTGGCCCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
Db 241 CCGAGGGGGTCTGGCCCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 346 CCGCTTCAATGACCTCAACCGGAGAGCTGTGAAATAGGGCTTTCCGAGTGGCATCTCG 405
Db 301 CCGCTTCAATGACCTCAACCGGAGAGCTGTGAAATAGGGCTTTCCGAGTGGCATCTCG 360
QY 406 CAACCAATGCTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
Db 361 CAACCAATGCTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 466 TGTTCGTGAGCTCTCTGAGTGGCTTTGTCTAAGCTGTGAGAGAGAGAGAGAGAG 525
Db 421 TGTTCGTGAGCTCTCTGAGTGGCTTTGTCTAAGCTGTGAGAGAGAGAGAGAGAG 480
QY 526 ACCTCTGAG 585
Db 481 ACCTCTGAG 540
QY 586 AGCAGAGATATTTGAG 645
Db 541 AGCAGAGATATTTGAG 600
QY 646 CAGGTGTGTCTCCCTTTGTGTAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 703
Db 601 CAGGTGTGTCTCCCTTTGTGTAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
```





|    |     |                                                                 |     |
|----|-----|-----------------------------------------------------------------|-----|
| Qy | 401 | CTCGGGAACCAATGCTGTGGAGCGGTGACCTCCATCCGCGCTCTTCCGCTCATGATG       | 460 |
|    |     |                                                                 |     |
|    |     |                                                                 |     |
| Db | 241 | CTTCGGGAACCAATGCTGTGGAGCGGTGACCTCCATCCGCGCTCTTCCGCTCATGATG      | 300 |
| Qy | 461 | CTTGAGTTCGATGAGCCCTCCTCTGGATTGAGCCCTGCTCATCGATGTCACCACTGAATCAG  | 520 |
|    |     |                                                                 |     |
|    |     |                                                                 |     |
| Db | 301 | CTTGAGTTCGATGAGCCCTCCTCCTGTGATGACCTTGTCTACCTGATGTCACCACTGAATCAG | 360 |
| Qy | 521 | CGGTG                                                           | 525 |
|    |     |                                                                 |     |
| Db | 361 | CGGTG                                                           | 365 |

```

RESULT 9
US-10-076-747-64
; Sequence 64, Application US/10076747
; Publication No. US20030180726A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and F
; FILE REFERENCE: DEX-0315
; CURRENT APPLICATION NUMBER: US/10/076,747
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,290
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/268,834
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-076-747-64

```

[illegible]

Db 450 SCATTCCTAATTATGTGGGACACCTTTCTCACTAATAATATGATATATTACGCGCAAAAAAAAA 509

Qy 1071 AAAAAAAAAAAAAAAAAAGCTGTGGGGGCTTACCCCGGGGCCAAGGGGGGGGCCCGGGGGGAAAT 1120

Db 510 AAAAAAAAAAAAAAAAAAGCGCTGGGGGGAAACCGGGGCCAAGCGCTGTTCGCGGGGTGACAT 569

Qy 1131 GTGTTCTCCCGGCCCAATTCCTCCCAATTTTGCATA 1170

Db 570 TGGTTTCCCGGCCCAATTTTCCACAAATTTGGGACAAA 609

RESULT 10  
US-10-027-632-174763/c  
; Sequence 174763, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:

```

1 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
2 TITLE OF INVENTION: Polymorphisms in the Human Genome
3 FILE REFERENCE: 108827.129
4 CURRENT APPLICATION NUMBER: US/10/027,632
5 CURRENT FILING DATE: 2002-04-30
6 PRIOR APPLICATION NUMBER: US 60/218, 006
7 PRIOR FILING DATE: 2000-07-12
8 PRIOR APPLICATION NUMBER: US 60/198, 676
9 PRIOR FILING DATE: 2000-04-20
10 PRIOR APPLICATION NUMBER: US 60/193, 463
11 PRIOR FILING DATE: 2000-03-29
12 PRIOR APPLICATION NUMBER: US 60/185, 218
13 PRIOR FILING DATE: 2000-02-24
14 PRIOR APPLICATION NUMBER: US 60/167, 363
15 PRIOR FILING DATE: 1999-11-23
16 PRIOR APPLICATION NUMBER: US 60/156, 358
17 PRIOR FILING DATE: 1999-09-28
18 PRIOR APPLICATION NUMBER: US 60/146, 002
19 PRIOR FILING DATE: 1999-08-09
20 NUMBER OF SEQ. ID NOS: 325720
21 SOFTWARE: FastSeq for Windows Version 4.0
22 SEQ ID NO 174763
23 LENGTH: 2940917
24 TYPE: DNA
25 ORGANISM: Human
26 FEATURE:
27 NAME/KEY: misc_feature
28 LOCATION: (1)..(2940917)
29 OTHER INFORMATION: n = A,T,C or G
30 US-10-027-632-174763

```

| Query Match           | 9.4%;        | Score 112.4;                                                 | DB 13;         | Length 2940917; |
|-----------------------|--------------|--------------------------------------------------------------|----------------|-----------------|
| Best Local Similarity | 66.5%        | Pred. No. 4.3e-17;                                           |                |                 |
| Matches 161;          | Conservative | 0;                                                           | Mismatches 81; | Indels 0;       |
|                       |              |                                                              |                | Gaps 0;         |
| Qy                    | 840          | TCACGTGATGCCAACATGTCGCATCCCTGTGGAGACATGCCGACGCTCTCTGCT       | 899            |                 |
| Db                    | 2551564      | TCCTTCGATACCAACATTTTCCGATACCTGTGTGCACATGTGGGGCTCGAATTGCGT    | 25515050       |                 |
| Qy                    | 900          | AGGAGACAGACCTTGCGCTTAGGGCAGGTGGAAAAATTCAGACTTTTTAGACACTGT    | 959            |                 |
| Db                    | 2551504      | AGGAGACCGAGATGACGAGGGGAGGCTTAAAGATTCCTAAACCTTTTAGCATGTGTT    | 2551445        |                 |
| Qy                    | 960          | TTTGTTTAATGGTATTTTTTATTGGCTACTTATTTGTTTAGCAAGTGTAGTGGCA      | 1019           |                 |
| Db                    | 2551444      | TTTACCTGTATGGACGCTTTTGTGCTTCAGATTACTTTGGGACATGTCGTGACAA      | 2551385        |                 |
| Qy                    | 1020         | TTCTATTATTGTGACCTTTTCATATAATGATTTAAGTAAAAA                   | 1079           |                 |
| Db                    | 2551384      | TGTTAGATTTTATTGTCTTTTATTATAGTATTTTAAAGTAAAGATGCGTCGATTAAGAGA | 2551325        |                 |
| Qy                    | 1080         | AA 1081                                                      |                |                 |
| Db                    | 2551324      | AA 2551323                                                   |                |                 |

```
RESULT 11
US-10-027-632-174763/c
; Sequence 174763, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174763
; LENGTH: 2940917
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2940917)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174763

Query Match 9.4%; Score 112.4; DB 16; Length 2940917;
Best Local Similarity 66.5%; Pred. No. 4.3e-17;
Matches 161; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 840 TCAGCTGATGCCAATGTTCCGATGCTGTCGAGATGCCGACGCTCTCCCTGCT 899
DB 2551564 TCTTTCATACCAACATTTTCCGATACCTGTTGCGATGCGCGCTGAGTTGCT 2551505
QY 900 AGGAGCAGAGACTTGGGCTTACGGGAGGAGTGAAGAAAATTCAGACTTTTTCAGCACTGTT 959
DB 2551504 AGGTGACCGGAGTATCGGAGGGGAGGCTAAAGAAATTCAGACTTTTTCAGCACTGTT 2551445
QY 960 TTGCTTTAATGCTAATTTTATTTGCTTACTTTATTTAGGACAAAGTGTAGTGGCA 1019
DB 2551444 TTACCTGATGAGAGGTTTGTGCTTACGATTAATCTTTGGGACTTACGCTGGACA 2551385
QY 1020 TTCTATTATTTGACCTTTTCATTAATPAGATTATTAAGTAAAGAAAAA 1079
DB 2551384 TGTTAAGTTTATTTGCTTTTATTTAAGTAATTTTAAGTAAGATCCGCTGATTAAGACA 2551325
QY 1080 AA 1081
DB 2551324 AA 2551323

RESULT 12
US-10-074-475-38
; Sequence 38, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Kaira, Kalpana
; APPLICANT: Cafferkey, Robert
```

```
APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-38

Query Match 6.3%; Score 75.6; DB 15; Length 461;
Best Local Similarity 62.9%; Pred. No. 1.2e-09;
Matches 117; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 986 GCTACTTATTTAGTACAGAGTGTAGTGCATTCATTTATTTGACCTTTCAATA 1045
DB 271 GCTTTTCTTTTGTGTGCTCAAGTGTAGCTGCATTTTAAATTTTGTGAATAAATTAAC 330
QY 1046 AATGATTTAATGAAAAAAGCTGTGAGGCTTACCGGCGG 1105
DB 331 TTGTGATTAAGAAAAAAGCTGTGAGGCTTACCGGCGG 390
QY 1106 CCAAAAGGGGCGGCGGAGATGTGCTTCCGCCCAATTCGCCCAATTTTGG 1165
DB 391 CCAAAAGGGGCGGCGGAGATGTGCTTCCGCCCAATTCGCCCAATTTTGG 450
QY 1166 AAAAA 1171
DB 451 AAAAA 456

RESULT 13
US-10-074-475-39
; Sequence 39, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Kaira, Kalpana
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-39

Query Match 6.3%; Score 75.6; DB 15; Length 633;
Best Local Similarity 62.9%; Pred. No. 1.5e-09;
Matches 117; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 986 GCTACTTATTTAGTACAGAGTGTAGTGCATTCATTTATTTGACCTTTCAATA 1045
DB 443 GCTTTTCTTTTGTGTGCTCAAGTGTAGCTGCATTTTAAATTTTGTGAATAAATTAAC 502
```

Tue Aug 24 09:48:34 2004

us-10-001-885-31.rnpb

Page 10

QY 1046 AATGATTATTAAGTAAAAAAGCTGTGGGTTACCCGGG 1105  
DB 503 TTTGTATTAGAAAAAAGCTGTGGGTTACCCGGG 562  
QY 1106 CCAAGGGGGGCGGCGGGAATGTCTCCGCCCAATTCGCCCAATTTTG 1165  
DB 563 CCAAGGGGGGCGGCGGGAATGTCTCCGCCCAATTCGCCCAATTTTG 622  
QY 1166 ACAGAA 1171  
DB 623 AGAGAG 628

RESULT 14

US-10-074-475-127  
Sequence 127, Application US/10074475  
Publication No. US2003092898A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Hu, Ping  
APPLICANT: Recipon, Hervé  
APPLICANT: Katta, Kalpana  
APPLICANT: Cafferkey, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific  
TITLE OF INVENTION: Genes and Proteins  
FILE REFERENCE: DEX-0313  
CURRENT APPLICATION NUMBER: US/10/074,475  
CURRENT FILING DATE: 2002-02-13  
PRIOR APPLICATION NUMBER: 60/268,292  
PRIOR FILING DATE: 2001-02-13  
NUMBER OF SEQ ID NOS: 295  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 127  
LENGTH: 420  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-074-475-127

Query Match 6.0%; Score 71.2; DB 15; Length 420;  
Best Local Similarity 75.9%; Pred. No. 1.8e-08;  
Matches 88; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1055 AAGTAAAAAAGCTGTGGGTTACCCGGGCGCAAGGGG 1114  
DB 302 AAAAAAAGCTGTGGGTTACCCGGGCGCAAGGGG 361  
QY 1115 GGGCGGGGGGGAATGTCTCCGCCCAATTCGCCCAATTTTGACAA 1170  
DB 362 GGGCGGGGGGGAATGTCTCCGCCCAATTCGCCCAATTTTGACAA 417

RESULT 15

US-10-021-323-6568  
Sequence 6568, Application US/10021323  
Publication No. US20040123340A1  
GENERAL INFORMATION:  
APPLICANT: Deikman, Jill  
APPLICANT: Peng, Paul C.C.  
APPLICANT: Fincher, Karen L.  
APPLICANT: Ziegler, Todd E.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21 (52274) B  
CURRENT APPLICATION NUMBER: US/10/021,323  
CURRENT FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: US 60/255, 619  
PRIOR FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 17880  
SEQ ID NO 6568

LENGTH: 480  
TYPE: DNA  
ORGANISM: Gossypium hirsutum  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(480)  
OTHER INFORMATION: unsure at all n locations  
OTHER INFORMATION: Clone ID: LIB828-008-Q1-N6-A12  
US-10-021-323-6568

Query Match 6.0%; Score 71.2; DB 17; Length 480;  
Best Local Similarity 62.2%; Pred. No. 1.9e-08;  
Matches 112; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 945 TTTTGTGACGTTTGTGTTAATGTAATTTTATGCTACTTATGTTAGGA 1004  
DB 21 TTTTGTGACGTTTGTGTTAATGTAATTTTATGCTACTTATGTTAGGA 80  
QY 1005 CAAGTGTAGTGCATTTATTTGACCTTTCAATTAATTAAGTAAAGGA 1064  
DB 81 TTTTGTGACGTTTGTGTTAATGTAATTTTATGCTACTTATGTTAGGA 140  
QY 1065 AAAAAAAGCTGTGGGTTACCCGGGCGCAAGGGGCGCGCGGG 1124  
DB 141 AAAAAAAGCTGTGGGTTACCCGGGCGCAAGGGGCGCGCGGG 200

Search completed: August 22, 2004, 04:52:06  
Job time : 605 secs

SUMMARIES

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 03:31:33 ; Search time 4687 Seconds  
(without alignments)  
11023.020 Million cell updates/sec

Title: US-10-001-885-31  
Perfect score: 1192  
Sequence: 1 aaacgcctcgagccgcgcg.....gaagagacacacaggaacc 1192

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_da:\*

2: gb\_hcg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pac:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_str:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_da:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pac:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_str:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_hcg\_hum:\*

31: em\_hcg\_inv:\*

32: em\_hcg\_other:\*

33: em\_hcg\_mus:\*

34: em\_hcg\_pln:\*

35: em\_hcg\_rod:\*

36: em\_hcg\_man:\*

37: em\_hcg\_vrl:\*

38: em\_sy:\*

39: em\_hlgo\_hum:\*

40: em\_hlgo\_mus:\*

41: em\_hlgo\_other:\*

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 1192  | 100.0       | 1192   | 6     | AX431088 Sequence   |
| 2          | 1033  | 86.7        | 1129   | 6     | BC013587 Homo sapi  |
| 3          | 792   | 66.4        | 1049   | 6     | BD191093 186 human  |
| 4          | 693   | 58.1        | 165110 | 9     | AL450311 Human DNA  |
| 5          | 689   | 57.8        | 173341 | 2     | AC021954 Homo sapi  |
| 6          | 473   | 39.7        | 483    | 6     | AR413227 Sequence   |
| 7          | 473   | 39.7        | 483    | 6     | BD108780 EST and e  |
| 8          | 369   | 31.0        | 550    | 6     | AX113082 Sequence   |
| 9          | 311   | 26.1        | 480    | 11    | G22561 human STR W  |
| 10         | 185   | 15.5        | 519    | 6     | AR415427 Sequence   |
| 11         | 185   | 15.5        | 519    | 6     | BD110980 EST and e  |
| 12         | 125   | 10.5        | 643    | 6     | AX431087 Sequence   |
| 13         | 38    | 3.2         | 68584  | 9     | AC019069 Homo sapi  |
| 14         | 38    | 3.2         | 118521 | 9     | AC119049 Homo sapi  |
| 15         | 38    | 3.2         | 118521 | 9     | HMAC004514          |
| 16         | 38    | 3.2         | 118521 | 9     | AC040159 Homo sapi  |
| 17         | 38    | 3.2         | 133456 | 9     | AC009111 Homo sapi  |
| 18         | 37    | 3.1         | 193014 | 2     | AC139357 Mus muscu  |
| 19         | 36    | 3.0         | 201371 | 2     | AC119162 Homo sapi  |
| 20         | 36    | 3.0         | 242372 | 2     | BX571714 Dantio rer |
| 21         | 36    | 3.0         | 328561 | 2     | BX321879 Dantio rer |
| 22         | 35    | 2.9         | 3255   | 3     | AY089465 Drosophill |
| 23         | 35    | 2.9         | 148955 | 5     | BX004813 Zebrafish  |
| 24         | 35    | 2.9         | 155287 | 2     | BX284669 Dantio rer |
| 25         | 35    | 2.9         | 189180 | 2     | AC023979 Homo sapi  |
| 26         | 35    | 2.9         | 200041 | 2     | BX571972 Dantio rer |
| 27         | 35    | 2.9         | 209817 | 10    | AL627385 Mouse DNA  |
| 28         | 35    | 2.9         | 232948 | 2     | AC107741 Mus muscu  |
| 29         | 34    | 2.9         | 436    | 9     | BC062727 Homo sapi  |
| 30         | 34    | 2.9         | 505    | 3     | AF379640 Heliothis  |
| 31         | 34    | 2.9         | 604    | 10    | BC025044 Mus muscu  |
| 32         | 34    | 2.9         | 1287   | 9     | BC009221 Homo sapi  |
| 33         | 34    | 2.9         | 1836   | 9     | BC046243 Homo sapi  |
| 34         | 34    | 2.9         | 2232   | 9     | AK130049 Homo sapi  |
| 35         | 34    | 2.9         | 2836   | 6     | AR170129 Sequence   |
| 36         | 34    | 2.9         | 2836   | 6     | AR170130 Sequence   |
| 37         | 34    | 2.9         | 2836   | 6     | BD069549 Novel car  |
| 38         | 34    | 2.9         | 2836   | 6     | BD069550 Novel car  |
| 39         | 34    | 2.9         | 3010   | 9     | BC060865 Homo sapi  |
| 40         | 34    | 2.9         | 47385  | 2     | AC104240 Homo sapi  |
| 41         | 34    | 2.9         | 81806  | 10    | AF326737 Mus muscu  |
| 42         | 34    | 2.9         | 109612 | 9     | AL512503 Human DNA  |
| 43         | 34    | 2.9         | 119456 | 2     | AC084789 Homo sapi  |
| 44         | 34    | 2.9         | 122619 | 10    | AL663050 Mouse DNA  |
| 45         | 34    | 2.9         | 123585 | 9     | AC004707 Homo sapi  |

ALIGNMENTS

RESULT 1  
AX431088  
LOCUS AX431088 1192 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 31 from Patent WO0240535.  
ACCESSION AX431088  
VERSION AX431088.1 GI:21655665  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 Recipon,H., Liu,C., Salceda,S., Sun,Y., Cafferkey,R. and  
Maccina,R.A.  
TITLE Compositions and methods relating to ovary specific genes and

Pred. No. is the number of results predicted by chance to have a

proteins  
Patent: WO 0240535-A 31 MAY-2002;  
DIADEXUS INC (US)  
Location/Qualifiers  
1. 1192  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 1192; DB 6; Length 1192;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AAACCCGCTGAGAGCCGCGGAGATGAGACGCCGCCGAGAGCCCGGAGATGCGCTGAGACA 60
DB 1 AAACCCGCTGAGAGCCGCGGAGATGAGACGCCGCCGAGAGCCCGGAGATGCGCTGAGACA 60
QY 61 CAGCATCTACAGCGTGGTCACTCTGTAAATCATCATGATGCAAGCCCTCCCTCAAT 120
DB 61 CAGCATCTACAGCGTGGTCACTCTGTAAATCATCATGATGCAAGCCCTCCCTCAAT 120
QY 121 TTCTGTGACAGCCCATGAGGAGCCACAGCGCTGAGAGATGATGCGATCTTGGCCAA 180
DB 121 TTCTGTGACAGCCCATGAGGAGCCACAGCGCTGAGAGATGATGCGATCTTGGCCAA 180
QY 181 TGGGGAATGTGAGAGACGACGACGCCCGAGATGAGAGCACTACCAAGCACCAAGAG 240
DB 181 TGGGGAATGTGAGAGACGACGACGCCCGAGATGAGAGCACTACCAAGCACCAAGAG 240
QY 241 TAGCATCTCTGACAGAGCTTCTTCATAGAGGCGCATGATGCTCCCGCAGGGGGCTCG 300
DB 241 TAGCATCTCTGACAGAGCTTCTTCATAGAGGCGCATGATGCTCCCGCAGGGGGCTCG 300
QY 301 CCCCCGACAG 360
DB 301 CCCCCGACAG 360
QY 361 CAACCGGACAGCTGTGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 CAACCGGACAGCTGTGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 GCGGATGACCTGACAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
DB 421 GCGGATGACCTGACAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 481 CCGTGTGAG 540
DB 481 CCGTGTGAG 540
QY 541 TAGGAG 600
DB 541 TAGGAG 600
QY 601 AGGAGATCTGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 AGGAGATCTGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 CCTTGTGTTAAG 720
DB 661 CCTTGTGTTAAG 720
QY 721 GGTTCGTTGTTCCGAGATGATCTTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 721 GGTTCGTTGTTCCGAGATGATCTTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 781 TCGAG 840
DB 781 TCGAG 840
QY 841 CAGGTGATGACCAATGTTCCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 841 CAGGTGATGACCAATGTTCCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
```

```
QY 901 GGGAGACAGACTTGGGCTTAGGGCAGGTGAGAAAAAATTCAGACTTTTGTAGCATGTGTT 960
DB 901 GGGAGACAGACTTGGGCTTAGGGCAGGTGAGAAAAAATTCAGACTTTTGTAGCATGTGTT 960
QY 961 TTGTTTATGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1020
DB 961 TTGTTTATGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1020
QY 1021 TCTATTATTTGATCTTTTCAATTAATAGATTTAAGTAAAAAATTTAAAAAATTTAAAA 1080
DB 1021 TCTATTATTTGATCTTTTCAATTAATAGATTTAAGTAAAAAATTTAAAAAATTTAAAA 1080
QY 1081 AAAAAGCTGTGGGGGTTACCGCGGGCCAAAGGGGGGCCCGGGGGGAGATGTTCTCC 1140
DB 1081 AAAAAGCTGTGGGGGTTACCGCGGGCCAAAGGGGGGCCCGGGGGGAGATGTTCTCC 1140
QY 1141 CGCCCCAATTCGCCCAATTTTGAACAAATGAAGAGACACAGAACCC 1192
DB 1141 CGCCCCAATTCGCCCAATTTTGAACAAATGAAGAGACACAGAACCC 1192
```

## RESULT 2

BC013587  
LOCUS  
DEFINITION  
Homo sapiens chromosome 10 open reading frame 35, mRNA (cDNA clone

MG:9596 IMAGE:389656), complete cds.

BC013587  
VERSION  
BC013587.1 GI:15488915

## KEYWORDS

MG.  
Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

## AUTHORS

Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hale, F.,  
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Umed, T.B., Toshiki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loughran, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, X., Gibbs, R.A.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahney, J., Hellon, E., Kettman, M., Madan, A., Young, A.C., Shcherbina, Y.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherbina, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Buetow, K.H., Kravitz, M.I., Skalska, U., Smalins, D.E.,  
Scherer, A., Schein, J.E., Jones, S.J., and Marz, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

## JOURNAL

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REMARK

## COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [gcgaps-remail.nih.gov](mailto:gcgaps-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing Arr: Sequencing Group at the Stanford Human Genome

Clome distribution: MGC clome distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
Series: IRAK Plate: 14 Row: 1 Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gt.: 21687152.  
Location/Name: ttf:stf

**Source**

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:9596 IMAGE:389655"
/tissue_type="Pancreas, epitheloid carcinoma"
/clone_Tip="NH_MGC_70"

```

**CDS**

```

/gene=C160C13
/db_xref="LocusID:219738"
160..525
/codon_start=1
/product="chromosome 10 open reading frame 35"
/protein_id="AAH3587.1"
/db_xref="GI:15488920"
/db_xref="LocusID:219738"
/translation="MWTILANGITVDDDDPRVTTQPPRGSIQPSFFNGHAPRC
GRGQQQAGARQAGAPFPNDINRQLVNNGPQMTLGNHAEVPTISILLILFLMMTC
RGGLLGVYLVSHLSQR"

```

## ORIGIN

|                            |        |               |           |              |
|----------------------------|--------|---------------|-----------|--------------|
| Query Match                | 86.7%  | Score 1033;   | DB 9;     | Length 1129; |
| Best Local Similarity      | 99.9%; | Pred. No. 0;  |           |              |
| Matches 1083; Conservative | 0;     | Mismatches 1; | Indels 0; | Gaps 0;      |

|    |     |                                                                  |     |
|----|-----|------------------------------------------------------------------|-----|
| QY | 2   | AACCGCCCTGAGAGCCGCCGAGAGTGGACGCGCCGAGAGCCCGAGAGCTGGGCTCGACAGAC   | 61  |
| Db | 1   | AACCGCCCTGAGAGCCGCCGAGAGTGGACGCGCCGAGAGCCCGAGAGCTGGGCTCGACAGAC   | 60  |
| QY | 62  | AGCATCTACTGACGCGTGGGTCACTCTGTGAACATCACTGACTGCAAGCTCCTCAATT       | 121 |
| Db | 61  | AGCATCTACTGACGCGTGGGTCACTCTGTGAACATCACTGACTGCAAGCTCCTCAATT       | 120 |
| QY | 122 | TCTGTGTGAGCCCATCAGGAGCCCAACAGGGCTGTGGAGAGATGTCGCGATCTTGGCCAT     | 181 |
| Db | 121 | TCTGTGTGAGCCCATCAGGAGCCCAACAGGGCTGTGGAGAGATGTCGCGATCTTGGCCAT     | 180 |
| QY | 182 | GGGGAAATCGTGCAGAGACGACGACCCCGAGTGAAGACCACTACCCAGCCACCAAGAGT      | 241 |
| Db | 181 | GGGGAAATCGTGCAGAGATGACGACCCCGAGTGAAGACCACTACCCAGCCACCAAGAGT      | 240 |
| QY | 242 | AGCATTCCTCGACAGAGCTTCTCAATPAGGGGCCCATPGGGCTCCCCCAGGGGGTCTGGC     | 301 |
| Db | 241 | AGCATTCCTCGACAGAGCTTCTCAATPAGGGGCCCATPGGGCTCCCCCAGGGGGTCTGGC     | 300 |
| QY | 302 | CCCCGCCAGCAGCAGGAGAGTGCACAGGCTGGGTGCTGCTCAGTCCCCCTTCATAGACTTC    | 361 |
| Db | 301 | CCCCGCCAGCAGCAGGAGAGTGCACAGGCTGGGTGCTGCTCAGTCCCCCTTCATAGACTTC    | 360 |
| QY | 362 | AACCGGCAGCTGTGTAACAATGGGCTTTTCGACATGGGCATCTGGCAACCATCTGTGGAG     | 421 |
| Db | 361 | AACCGGCAGCTGTGTAACAATGGGCTTTTCGACATGGGCATCTGGCAACCATCTGTGGAG     | 420 |
| QY | 422 | COGGTGAACCTCCATCCTGTCTCTCTTCCTCATGATGATCTTGTGTTTCGTGGCTCTCTC     | 481 |
| Db | 421 | COGGTGAACCTCCATCCTGTCTCTCTTCCTCATGATGATCTTGTGTTTCGTGGCTCTCTC     | 480 |
| QY | 482 | CTGGTGTGACCTTGTGTACCTGGTGTGCCACCTGATTCAGAGCGGTGACACTGTAGAGGCTGAT | 541 |
| Db | 481 | CTGGTGTGACCTTGTGTACCTGGTGTGCCACCTGATTCAGAGCGGTGACACTGTAGAGGCTGAT | 540 |

|    |      |                        |                                           |                 |     |
|----|------|------------------------|-------------------------------------------|-----------------|-----|
| Qy | 542  | AGGGGTGGATTGTTGAGAGGGA | CTTGCTGGGCTTGTTGATGAGAGAGGATATTTGGA       | 601             |     |
| Db | 541  | AGGGTGGGTTTGTGAGAGGGAC | CTTGCTGGGCTTGTTGATGAGAGAGGATATTTGGA       | 600             |     |
| Qy | 602  | GGGGATCTGCTGCTGCTTGAAG | TATGATCAGAGAGGGACCA                       | CAGCTGTGTCTTCCC | 661 |
| Db | 601  | GGGGATCTGCTGCTGCTTGAAG | TATGATCAGAGAGGGACCA                       | CAGCTGTGTCTTCCC | 660 |
| Qy | 662  | CTTTGCTTTAAGCTGAGGAGG  | AGAGGAGCTTAAGTCACAGATTTCCAAAGTGTGGGTGG    | 721             |     |
| Db | 661  | CTTTGCTTTAAGCTGAGGAGG  | AGAGGAGCTTAAGTCACAGATTTCCAAAGTGTGGGTGG    | 720             |     |
| Qy | 722  | GTCCGTTGATTCCCGAATATCT | TTTAGTGGTATAGGGGCTGCATTAAGTGACACAAAT      | 781             |     |
| Db | 721  | GTCCGTTGATTCCCGAATATCT | TTTAGTGGTATAGGGGCTGCATTAAGTGACACAAAT      | 780             |     |
| Qy | 782  | CAGAGCAGAAAGGATGCCCTTC | CCAAATTCCTCAATCCTTTTATGCCGAGAAAGATCTC     | 841             |     |
| Db | 781  | CAGAGCAGAAAGGATGCCCTTC | CCAAATTCCTCAATCCTTTTATGCCGAGAAAGATCTC     | 840             |     |
| Qy | 842  | AGCTGATGCCCAATGTTCCGAT | GCCTGTGGAAAGATCCGACGCTCCTCTGSCCTAG        | 901             |     |
| Db | 841  | AGCTGATGCCCAATGTTCCGAT | GCCTGTGGAAAGATCCGACGCTCCTCTGSCCTAG        | 900             |     |
| Qy | 902  | GGAGCAGAGACTTGGGGCTTAG | GGGACAGTGGAAAAATTCAGACTTTTATAGCATGTGTTT   | 961             |     |
| Db | 901  | GGAGCAGAGACTTGGGGCTTAG | GGGACAGTGGAAAAATTCAGACTTTTATAGCATGTGTTT   | 960             |     |
| Qy | 962  | TGTTTTAATGATATATTTTAT  | TGTTATGCTACTTATTTGTTTAGACAAAGTGTAGTGCCATT | 1021            |     |
| Db | 961  | TGTTTTAATGATATATTTTAT  | TGTTATGCTACTTATTTGTTTAGACAAAGTGTAGTGCCATT | 1020            |     |
| Qy | 1022 | CTATTTATGTGACCTTTCAAT  | ATAATAGATTTAAGTMAAAAAAAAAAAAAAAAAAAAAA    | 1081            |     |
| Db | 1021 | CTATTTATGTGACCTTTCAAT  | ATAATAGATTTAAGTMAAAAAAAAAAAAAAAAAAAAAA    | 1080            |     |
| Qy | 1082 | AAAA                   | 1085                                      |                 |     |
| Db | 1081 | AAAA                   | 1084                                      |                 |     |

RESULT 3  
BD191093

|            |                              |             |     |        |                 |
|------------|------------------------------|-------------|-----|--------|-----------------|
| LOCUS      | BD191093                     | 1049 bp     | DNA | linear | PAT 17-JUL-2003 |
| DEFINITION | 186 human secreted proteins. |             |     |        |                 |
| ACCESSION  | BD191093                     |             |     |        |                 |
| VERSION    | BD191093.1                   | GI:33000832 |     |        |                 |
| KEYWORDS   | JP 2002510192-A/57.          |             |     |        |                 |
| SOURCE     | unidentified                 |             |     |        |                 |

## REFERENCE

**AUTHORS** Kudenov, S.M., Kosen, C.A., Fischer, C.L., Soppet, J.R., Carter, A.C., Bednatsk, D.P., Endress, G.A., Yu, G.L., N.J., Feng, P., Young, P.E., Greens, J.M., Ferris, A.M., Dunn, R., Hu, J.S., Florence, K.A., Olsen, H.S., Ebner, R., Brewer, L.A., Moore, P.A., Shi, Y., Lafleur, D.W., Li, Y., Zeng, Z. and Kyaw, H.

**TITLE** 186 human secreted proteins

**JOURNAL** Patent: JP 2002510192-A-57 02-APR-2002;

COMMENT

PD 02-APR-2002  
PF 06-MAR-1998 JP 1998538883  
PR 07-MAR-1997 US 60/040162, 07-MAR-1997 US 60/040333 PR  
07-MAR-1997 US 60/038621, 07-MAR-1997 US 60/040161  
07-MAR-1997 US 60/040626, 07-MAR-1997 US 60/040334 PR  
07-MAR-1997 US 60/040366, 07-MAR-1997 US 60/040363 PR  
11-APR-1997 US 60/043560, 11-APR-1997 US 60/043568 PI  
M RUBEN, CRAIG A ROSEN, CARRIE L FISCHER, DANIEL R SOPPET, PI  
KENNETH C CATER, DANIEL P BEDNARIK, GREGORY  
A ENDRSS, GUO LIANG  
PI YU, TIAN NT  
PI PING FENG, PAUL E YOUNG, JOHN M GREENE, ANN



M FERRIE, ROXANNE DUAN,  
PI JING SHAN HU, KIMBERLY A FLORENCE, HENRIK  
S OLSEN, REINHARD ERNER,  
PI LAURIE A BREWER, PAUL A MOORE, YANGSU SHI, DAVID W LAFLEUR PI  
, YI LI, ZHI ZHEN ZENG,  
PI HLA KYAM  
PC C12N15/12, C12N5/10, C12N1/21, C07K14/47, C07K16/18, C12Q1/68, PC  
G01N33/50,  
PC G01N33/53, G01N33/68, A61K38/17  
CC Strandness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
1. 1049  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

Query Match 66.4%; Score 792; DB 6; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

ORIGIN  
46 TCGGCGCTGCGAGACAGCATCTACTAGCGTGGGTCACTCTGTGAACATCACTGACTG 105  
1 TCGGCGCTGCGAGACAGCATCTACTAGCGTGGGTCACTCTGTGAACATCACTGACTG 60  
106 CAAGCTCTCCCTCAATTTCTGGTGAAGCCCATCAAGGACCCACAGCGCTGGAGATGGT 165  
61 CAAGCTCTCCCTCAATTTCTGGTGAAGCCCATCAAGGACCCACAGCGCTGGAGATGGT 120  
166 GCGGATCTTGGCCCAATGAGGAAATGTCGACAGACAGACAGACAGACAGACAGACAG 225  
121 GCGGATCTTGGCCCAATGAGGAAATGTCGACAGACAGACAGACAGACAGACAGACAG 180  
226 CCAGCCCAAGAGAGTGAATCTCTGACAGAGCTTCTCAATAGGAGGACATGATCTCC 285  
181 CCAGCCCAAGAGAGTGAATCTCTGACAGAGCTTCTCAATAGGAGGACATGATCTCC 240  
286 CCAGAGGAGTCTGAGCCCGCCGACAGACAGAGAGTCCAGAGCTGAGTCTGAGTCC 345  
241 CCAGAGGAGTCTGAGCCCGCCGACAGAGAGAGTCCAGAGCTGAGTCTGAGTCC 300  
346 CCCCTTCAATGACCTCAACCGGACGCTGAGAACATGAGCTTCCGACATGGCATCTCG 405  
301 CCCCTTCAATGACCTCAACCGGACGCTGAGAACATGAGCTTCCGACATGGCATCTCG 360  
406 CAACCATGCTGAGAGCGGAGCTGACCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTG 465  
361 CAACCATGCTGAGAGCGGAGCTGACCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTG 420  
466 TGTTCGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 525  
421 TGTTCGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
526 ACCCTTAAGGAGCTATAGGAGGAGTGTGTTGTAAGAGGAGTGTGTTGTAAGAGGAGT 585  
481 ACCCTTAAGGAGCTATAGGAGGAGTGTGTTGTAAGAGGAGTGTGTTGTAAGAGGAGT 540  
586 AGCAGGATATTTGAGAGGAGTGTGTTGTAAGAGGAGTGTGTTGTAAGAGGAGTGTG 645  
541 AGCAGGATATTTGAGAGGAGTGTGTTGTAAGAGGAGTGTGTTGTAAGAGGAGTGTG 600  
646 CAGGTGTGTTTCCCTTGTGTGTAAGAGGAGTGTGTTGTAAGAGGAGTGTGTTGTAAG 704  
601 CAGGTGTGTTTCCCTTGTGTGTAAGAGGAGTGTGTTGTAAGAGGAGTGTGTTGTAAG 660  
705 TT-CGAAAGTGTGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 763  
661 TTCCCAAGTGTGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720  
764 CATTAATGCTGCAATATATATATATATATATATATATATATATATATATATATATAT 823

Db 721 CATTAATGCTGCAATATATATATATATATATATATATATATATATATATATATAT 780  
Qy 824 TTATGCCGAGAAAGATCTGAGCTGATGCCACATGTTCCGATGCTGTGGAAGACATGCC 883  
Db 781 TTATGCCGAGAAAGATCTGAGCTGATGCCACATGTTCCGATGCTGTGGAAGACATGCC 840  
Qy 884 GACGTCTCTCTGCTTGAAGAGACAGACTTGGGCTTGAAGAGAGTGGAAAAATTCGAGA 943  
Db 841 GACGTCTCTCTGCTTGAAGAGACAGACTTGGGCTTGAAGAGAGTGGAAAAATTCGAGA 900  
Qy 944 CTTTATTAGCACTGTTTGTGTTTATGAGTATTTTATTTGCTACTTTATTTGTTAGG 1003  
Db 901 CTTTATTAGCACTGTTTGTGTTTATGAGTATTTTATTTGCTACTTTATTTGTTAGG 960  
Qy 1004 ACAAGTGTAGTGCATCTATTTATGTTGACCTTTCAATTAATTAATTAATTAATTA 1063  
Db 961 ACAAGTGTAGTGCATCTATTTATGTTGACCTTTCAATTAATTAATTAATTAATTA 1020  
Qy 1064 AAAAAAAAAAAAAA 1077  
Db 1021 AAAAAAAAAAAAAA 1034

RESULT 4  
AL450311  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AL450311 165110 bp DNA linear PRI 12-JUN-2001  
Human DNA sequence from clone RP11-343J3 on chromosome 10, complete  
sequence.  
AL450311  
AL450311.11 GI:14626972  
HTG.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 165110)  
Howden, P.  
Direct Submission  
Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requesters: clonerequest@sanger.ac.uk  
On Jul 8, 2001 this sequence version replaced gi:14575291.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em, EMBL; SW,  
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 10, constructed by the Sanger Centre Chromosome 10  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr10  
RP11-343J3 is from the library RPCR-11.2 constructed by the group  
of Pieper de Jong. For further details see  
http://www.choyl.org/bacpac/home.htm  
VECTOR: pBAC3.6  
This sequence is the entire insert of clone RP11-343J3. The true  
left end of clone RP11-242G20 is at 139955 in this sequence. The  
true right end of clone RP11-404C6 is at 6588 in this sequence.

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/clone="RP11-343J3"  
/clone\_1b="RP11-11.2"  
7.147  
repeat\_region /note="TH1C repeat: matches 2. .142 of consensus"  
151. .293  
repeat\_region /note="MIR repeat: matches 2. .148 of consensus"  
1136. .1187  
repeat\_region /note="26 copies 2 mer gt 98% conserved"  
2960. .3372  
repeat\_region /note="AluSg repeat: matches 1. .310 of consensus"  
4144. .4270  
repeat\_region /note="AluX repeat: matches 1. .134 of consensus"  
4319. .4494  
repeat\_region /note="AluX repeat: matches 118. .293 of consensus"  
5662. .6438  
repeat\_region /note="L1MC repeat: matches 272. .1095 of consensus"  
6916. .8080  
repeat\_region /note="L1MC repeat: matches 1168. .2367 of consensus"  
9934. .10087  
repeat\_region /note="MIR repeat: matches 25. .184 of consensus"  
10353. .10481  
repeat\_region /note="MIR repeat: matches 65. .194 of consensus"  
11077. .11372  
repeat\_region /note="AluX repeat: matches 1. .294 of consensus"  
11600. .11639  
repeat\_region /note="20 copies 2 mer tc 95% conserved"  
12479. .12608  
repeat\_region /note="65 copies 2 mer at 90% conserved"  
13433. .13876  
repeat\_region /note="MIR repeat: matches 1. .466 of consensus"  
13372. .14576  
repeat\_region /note="AluSg repeat: matches 1. .301 of consensus"  
14577. .14633  
repeat\_region /note="AluSg repeat: matches 243. .299 of consensus"  
14843. .14943  
repeat\_region /note="MER8 repeat: matches 9. .109 of consensus"  
15404. .15613  
repeat\_region /note="MIR repeat: matches 3. .213 of consensus"  
16626. .16705  
repeat\_region /note="MIR repeat: matches 60. .140 of consensus"  
16928. .16993  
repeat\_region /note="33 copies 2 mer gg 66% conserved"  
17476. .17559  
repeat\_region /note="MER81 repeat: matches 2. .114 of consensus"  
17719. .18069  
repeat\_region /note="L1MC4 repeat: matches 7617. .7977 of consensus"  
18117. .18345  
repeat\_region /note="MIR repeat: matches 7. .262 of consensus"  
19169. .19400  
repeat\_region /note="MIR repeat: matches 5. .239 of consensus"  
19436. .19627  
repeat\_region /note="L1MC5 repeat: matches 7728. .7917 of consensus"  
19642. .19716  
repeat\_region /note="MIR repeat: matches 48. .131 of consensus"  
19969. .20010  
repeat\_region /note="21 copies 2 mer tg 100% conserved"  
20623. .20704  
repeat\_region /note="41 copies 2 mer gt 85% conserved"  
20738. .20773  
repeat\_region /note="U2 repeat: matches 1. .36 of consensus"  
20802. .20863  
repeat\_region /note="MIR repeat: matches 1. .62 of consensus"  
20889. .21263  
repeat\_region /note="MIR repeat: matches 188. .541 of consensus"  
21463. .21618  
repeat\_region /note="MIR repeat: matches 46. .192 of consensus"  
22019. .22326  
repeat\_region /note="AluY repeat: matches 1. .306 of consensus"  
22381. .22564

/note="U2 repeat: matches 2453. .2629 of consensus"  
22896. .23174  
repeat\_region /note="MIR repeat: matches 117. .413 of consensus"  
23215. .23346  
repeat\_region /note="MIR repeat: matches 48. .188 of consensus"  
23388. .23532  
repeat\_region /note="U2 repeat: matches 2097. .2230 of consensus"  
23533. .23837  
repeat\_region /note="AluY repeat: matches 1. .303 of consensus"  
23838. .24137  
repeat\_region /note="U2 repeat: matches 1754. .2097 of consensus"  
24291. .24581  
repeat\_region /note="AluX repeat: matches 1. .300 of consensus"  
24653. .24850  
repeat\_region /note="MIR repeat: matches 1. .200 of consensus"  
26189. .26685  
misc\_feature /note="CpG island"  
29453. .30918  
/evidence=not\_experimental  
29453. .30918  
misc\_feature /note="CpG island"  
31084. .31156  
repeat\_region /evidence=not\_experimental  
31618. .31876  
repeat\_region /note="MIR repeat: matches 65. .138 of consensus"  
32767. .32830  
repeat\_region /note="AluY repeat: matches 29. .275 of consensus"  
33050. .33178  
repeat\_region /note="MIR repeat: matches 76. .139 of consensus"  
35112. .36201  
repeat\_region /note="43 copies 3 mer tcc 72% conserved"  
35112. .36201  
misc\_feature /note="CpG island"  
36551. .36604  
/evidence=not\_experimental  
36551. .36604  
repeat\_region /note="27 copies 2 mer ac 94% conserved"  
37585. .38254  
repeat\_region /note="L1P15 repeat: matches 5480. .6157 of consensus"  
38265. .38561  
repeat\_region /note="AluX repeat: matches 1. .300 of consensus"  
39214. .39423  
repeat\_region /note="MIR repeat: matches 3. .219 of consensus"  
39489. .39601  
repeat\_region /note="AluX repeat: matches 1. .292 of consensus"  
40050. .40189  
repeat\_region /note="MIR repeat: matches 1. .144 of consensus"  
41057. .41290  
repeat\_region /note="MIR repeat: matches 8. .240 of consensus"  
41296. .41460  
repeat\_region /note="U2 repeat: matches 2569. .2730 of consensus"  
41506. .41944  
repeat\_region /note="U2 repeat: matches 1916. .2416 of consensus"  
42388. .42698  
repeat\_region /note="U2 repeat: matches 1448. .1779 of consensus"  
44193. .44579  
repeat\_region /note="TH1C repeat: matches 1. .371 of consensus"  
44600. .44956  
repeat\_region /note="L1P16 repeat: matches 90. .445 of consensus"  
45240. .45300  
repeat\_region /note="MER58A repeat: matches 37. .97 of consensus"  
45798. .45909  
repeat\_region /note="L1P1 repeat: matches 90. .192 of consensus"  
46826. .46871  
repeat\_region /note="23 copies 2 mer gt 97% conserved"  
50684. .50969  
repeat\_region /note="AluY repeat: matches 1. .287 of consensus"  
50980. .51291  
repeat\_region /note="AluSg repeat: matches 1. .312 of consensus"  
52222. .52519  
repeat\_region /note="AluX repeat: matches 1. .300 of consensus"  
54065. .54260  
repeat\_region /note="L1P4 repeat: matches 3865. .4055 of consensus"  
54261. .54432  
repeat\_region /note="FAM repeat: matches 2. .167 of consensus"  
54433. .54629

/note="U1M4 repeat: matches 3652. .3865 of consensus"  
 54648. 54862  
 /note="U1R41 repeat: matches 11. .217 of consensus"  
 54863. 55236  
 /note="U1R1A1 repeat: matches 1. .365 of consensus"  
 55237. 55700  
 repeat\_region

Query Match 58.1%; Score 693; DB 9; Length 165110;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 793; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 269 AGGGGCGCATGGTGTCTCCCGAGGGGGTCTGGCCCGCCGAGCAGCAGCAGGTCGACG 328  
 DB 90084 AGGGGCGCATGGTGTCTCCCGAGGGGGTCTGGCCCGCCGAGCAGCAGCAGGTCGACG 90143  
 QY 329 CTGGGCTCTCTCAGTCTCCCGCTTCAATGACCTCAACCGAGCTGGTGAACATGGGCTTT 388  
 DB 90144 CTGGGCTCTCTCAGTCTCCCGCTTCAATGACCTCAACCGAGCTGGTGAACATGGGCTTT 90203  
 QY 389 CCGCAGTGGCATCTCGGCAACCATGCTGTGAGCCGGTGAACCTCCATCTGCTCTCTTC 448  
 DB 90204 CCGCAGTGGCATCTCGGCAACCATGCTGTGAGCCGGTGAACCTCCATCTGCTCTCTTC 90263  
 QY 449 CTGCTCATGATGCTTGGTGTCTGCTGCTCTCTCTCTGCTTGGCTTGTCTAAGTGTCTC 508  
 DB 90264 CTGCTCATGATGCTTGGTGTCTGCTGCTCTCTCTCTGCTTGGCTTGTCTAAGTGTCTC 90323  
 QY 509 CACCTGATGAGGGGTCACCTCGAGGGGTCATAGGGGGTGGGCTTTGTGAGAGGACCTTG 568  
 DB 90324 CACCTGATGAGGGGTCACCTCGAGGGGTCATAGGGGGTGGGCTTTGTGAGAGGACCTTG 90383  
 QY 569 CTGGGCTTGTGTGTGAGAGGAGCATATTGAGGGGATCTGTGTGCTCTTGAAGGAT 628  
 DB 90384 CTGGGCTTGTGTGTGAGAGGAGCATATTGAGGGGATCTGTGTGCTCTTGAAGGAT 90443  
 QY 629 GATAGAGAGGGGACCAAGGTGTGTTTCCCTTGTGTTAAGGTGAGGAGGAGGGA 688  
 DB 90444 GATAGAGAGGGGACCAAGGTGTGTTTCCCTTGTGTTAAGGTGAGGAGGAGGGA 90503  
 QY 689 GACGTTAGTCAGCATTTCCAAAGTGTGGGTGGTCCGTTGTTCCGAGATACCTTTAG 748  
 DB 90504 GACGTTAGTCAGCATTTCCAAAGTGTGGGTGGTCCGTTGTTCCGAGATACCTTTAG 90563  
 QY 749 GTGTTATGGGGCTCGCATTTAAGTGGCAAAATCAGAGCAAGAAAGCATGCCCTTCCA 808  
 DB 90564 GTGTTATGGGGCTCGCATTTAAGTGGCAAAATCAGAGCAAGAAAGCATGCCCTTCCA 90623  
 QY 809 ATTCTCTCANTCCTTTATGCGGAGAAATCTCAGCTGATGCGCAATGTTCCATGCC 868  
 DB 90624 ATTCTCTCANTCCTTTATGCGGAGAAATCTCAGCTGATGCGCAATGTTCCATGCC 90683  
 QY 869 TGTGAAAGCATGCGCAGTCTCTCTGCTTGAAGGAGCAGGACTTGGGCTTGAAGGAGGT 928  
 DB 90684 TGTGAAAGCATGCGCAGTCTCTCTGCTTGAAGGAGCAGGACTTGGGCTTGAAGGAGGT 90743  
 QY 929 GGAAGAAATTCGACGCTTTTATGACCTGTTTGTATGATGATTTATTTATGAGCT 988  
 DB 90744 GGAAGAAATTCGACGCTTTTATGACGCTGTTTGTATGATGATTTATTTATGAGCT 90803  
 QY 989 ACTTATTTGTTAGAGCAAGTGTGATGATTTATTTATTTGAGCCTTTTCAATTAAT 1048  
 DB 90804 ACTTATTTGTTAGAGCAAGTGTGATGATTTATTTATTTGAGCCTTTTCAATTAAT 90863  
 QY 1049 AGATTTAAGTAAAA 1063  
 DB 90864 AGATTTAAGTAAAA 90878

ACCESSION AC021954 GI:7417809  
 VERSION AC021954.3  
 KEYWORDS HTGS PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 173341)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens chromosome 10, clone R11-57E12  
 Unpublished  
 2 (bases 1 to 173341)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Beckert, R., Beda, F.,  
 Boguslavsky, L., Bouckgele, B., Brown, A., Burkett, G., Castle, A.,  
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 Dekrellano, K., Dewar, K., Domino, M., Doyle, M., Fenebor, J.,  
 Ferrerella, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,  
 Gardyna, S., Grant, G., Hagos, B., Heatord, A., Horton, L.,  
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Landers, T., Lehoczy, J., Levine, R., Liu, G., Locke, K.,  
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
 McNeelers, R., Meldrim, J., Menus, L., Morrow, J., Naylor, J.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,  
 Pletre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Subramanian, A., Talamas, J., Teefaye, S., Theodore, J.,  
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. D.,  
 Zimmer, A. and Zody, M.

TITLE JOURNAL  
 REFERENCE  
 AUTHORS  
 Submitted (22-JUN-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02144, USA  
 3 (bases 1 to 173341)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
 Boguslavsky, L., Bouckgele, B., Brown, A., Burkett, G.,  
 Campoliano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collymore, A., Cooke, P., Dekrellano, K., Dewar, K., Diaz, J. S.,  
 Dodge, S., Domino, M., Doyle, M., Ferrerella, P., Fitzhugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
 Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., McNeelers, R.,  
 McEwan, P., McGurk, A., McKernan, K., Menus, L.,  
 Meldrim, J., Menus, L., Milova, T., Miranda, C., Mlenga, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pletre, N.,  
 Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Teefaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. D.,  
 Young, G., Zainoun, D., Zimmer, A. and Zody, M.

TITLE JOURNAL  
 COMMENT  
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02144, USA  
 On Apr 5, 2000 this sequence version replaced gi:698455.  
 All repeats were identified using RepeatMasker:  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: W1BR  
 Web site: http://www-seg.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: U5931  
 Center clone name: 57\_E\_12  
 ----- Summary Statistics  
 Sequencing vector: M13; M7815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly: Phred version 0.960731

RESULT 5  
 AC021954/c 173341 bp DNA linear HNG 24-AUG-2002  
 LOCUS  
 DEFINITION Homo sapiens chromosome 10 clone R11-57E12 map 10, WORKING DRAFT

|                    |                                    |
|--------------------|------------------------------------|
| Consensus quality: | 161190 bases at least Q30          |
| Consensus quality: | 166837 bases at least Q30          |
| Consensus quality: | 169995 bases at least Q20          |
| Insert size:       | 176000; agarose-gel                |
| Insert size:       | 171041; sum-of-conflicts           |
| Quality coverage:  | 3.7 in Q20 bases; agarose-gel      |
| Quality coverage:  | 3.8 in Q20 bases; sum-of-conflicts |

|        |                                      |
|--------|--------------------------------------|
| 1      | 1574: contig of 1574 bp in length    |
| 1575   | 1674: gap of 100 bp                  |
| 1675   | 3403: contig of 1369 bp in length    |
| 3044   | 3143: gap of 100 bp                  |
| 3144   | 5889: contig of 2746 bp in length    |
| 5890   | 5959: gap of 100 bp                  |
| 5990   | 8979: contig of 2990 bp in length    |
| 8980   | 9079: gap of 100 bp                  |
| 9080   | 13649: contig of 4595 bp in length   |
| 13675  | 13774: gap of 100 bp                 |
| 13775  | 18831: contig of 5057 bp in length   |
| 18832  | 18931: gap of 100 bp                 |
| 18932  | 23526: contig of 4595 bp in length   |
| 23527  | 23626: gap of 100 bp                 |
| 23627  | 27386: contig of 3760 bp in length   |
| 27387  | 27486: gap of 100 bp                 |
| 27487  | 32572: contig of 5086 bp in length   |
| 32573  | 32672: gap of 100 bp                 |
| 32673  | 38632: contig of 5960 bp in length   |
| 38633  | 38732: gap of 100 bp                 |
| 38733  | 43725: contig of 5003 bp in length   |
| 43736  | 43835: gap of 100 bp                 |
| 43836  | 49020: contig of 5185 bp in length   |
| 49021  | 49120: gap of 100 bp                 |
| 49121  | 53600: contig of 4540 bp in length   |
| 53661  | 53760: gap of 100 bp                 |
| 53761  | 59544: contig of 5784 bp in length   |
| 59545  | 59644: gap of 100 bp                 |
| 59645  | 66335: contig of 6751 bp in length   |
| 66396  | 66435: gap of 100 bp                 |
| 66496  | 74606: contig of 8201 bp in length   |
| 74697  | 74796: gap of 100 bp                 |
| 74797  | 81336: contig of 6530 bp in length   |
| 81327  | 81426: gap of 100 bp                 |
| 81427  | 89446: contig of 8020 bp in length   |
| 89447  | 89546: gap of 100 bp                 |
| 89547  | 96459: contig of 6913 bp in length   |
| 96460  | 96559: gap of 100 bp                 |
| 96560  | 106488: contig of 9939 bp in length  |
| 106490 | 106558: gap of 100 bp                |
| 106599 | 116966: contig of 10386 bp in length |
| 116987 | 117086: gap of 100 bp                |
| 117087 | 128890: contig of 11804 bp in length |
| 128891 | 128990: gap of 100 bp                |
| 129391 | 147200: contig of 18300 bp in length |
| 147391 | 147330: gap of 100 bp                |
| 147391 | 173331: contig of 25951 bp in length |

```

misc_feature
 source
 location/Qualifiers
 1..173341
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /map="10"
 /clone="RP11-57E12"
 /clone_lib="RPC1-11 Human Male BAC"
 1..1574

```

```
misc_feature /note="assembly_fragment" 1675. .3043
misc_feature /note="assembly_fragment" 3144. .5589
misc_feature /note="assembly_fragment" 5990. .8979
misc_feature /note="assembly_fragment" 9080. .11674
misc_feature /note="assembly_fragment" 13775. 18831
misc_feature /note="assembly_fragment" 18932. .23526
misc_feature /note="assembly_fragment" 23627. .27386
misc_feature /note="assembly_fragment clone_end:SP6
vector_side:left" 27487. .32572
misc_feature /note="assembly_fragment" 32673. .38632
misc_feature /note="assembly_fragment" 38733. .43735
misc_feature /note="assembly_fragment" 43836. .49020
misc_feature /note="assembly_fragment" 49121. 53660
misc_feature /note="assembly_fragment" 53761. .59544
misc_feature /note="assembly_fragment" 59645. .66395
misc_feature /note="assembly_fragment" 66496. .74696
misc_feature /note="assembly_fragment" 74797. .81326
misc_feature /note="assembly_fragment" 81427. .89446
misc_feature /note="assembly_fragment" 89547. .96459
misc_feature /note="assembly_fragment" 96560. .106498
misc_feature /note="assembly_fragment" 106599. .116986
misc_feature /note="assembly_fragment" 117087. .128890
misc_feature /note="assembly_fragment clone_end:T7
vector_side:right" 128991. .147290
misc_feature /note="assembly_fragment" 147391. .173341
misc_feature /note="assembly_fragment"
```

|    | Query Match           | Similarity                                                     | Score        | DB 2:  | Length     | 173341;              |
|----|-----------------------|----------------------------------------------------------------|--------------|--------|------------|----------------------|
|    | Best Local Similarity | 99.9%                                                          | Pred.        | No. 0; |            |                      |
|    | Matches               | 809;                                                           | Conservative | 0;     | Mismatches | 0; Indels 1; Gaps 1; |
| QY | 254                   | CAGAGCTTTCATATGAGGGCCCATGTGTCCTCCCACAGGGGGCTCTGGCCCCCGCACAG    | 313          |        |            |                      |
| Db | 161089                | CAGAGCTTTCATATGAGGGCCCATGTGTCCTCCCACAGGGGGCTCTGGCCCCCGCACAG    | 161030       |        |            |                      |
| QY | 314                   | CAGGCAGGTCAGAGCTGGTGTCGTCAATGCCCTTCATATGACTCAACCCGACAGCTG      | 373          |        |            |                      |
| Db | 161029                | CAGGCAGGTCAGAGCTGGTGTCGTCAATGCCCTTCATATGACTCAACCCGACAGCTG      | 160970       |        |            |                      |
| QY | 374                   | GTTGAACATBGGCTTTCGCGAGTGGCATTCTCGGCCAACATGTGTGGAGCCGGTGAACCTCC | 423          |        |            |                      |
| Db | 160969                | GTTGAACATBGGCTTTCGCGAGTGGCATTCTCGGCCAACATGTGTGGAGCCGGTGAACCTCC | 160910       |        |            |                      |
| QY | 434                   | ATCCGAGCTCTCTTCTGCTCATGTATGTTGGTGTGTTGTGGACCTCCCTCGTTGGGCTT    | 493          |        |            |                      |
| Db | 160909                | ATCCGAGCTCTCTTCTGCTCATGTATGTTGGTGTGTTGTGGACCTCCCTCGTTGGGCTT    | 160850       |        |            |                      |
| QY | 494                   | GTTCACTGAGTGTCCCACTTAATGACGGTGAACCTCTGAAGGCTGATTAAGGGGTGGTTTT  | 553          |        |            |                      |

Db 160849 GTCTACCTGGTGTCCACCTGAGTCAAGCGGTGACCTCTGAGGCTGTATGAGGGTGGGGTTT 160790  
QY 554 GTTGAAGAGGAGCTTGTCTGGCTTGTGTGAGCAGCAGCATATTGGAGGGATCTGTG 613  
Db 160789 GTTGAAGAGGAGCTTGTCTGGCTTGTGTGAGCAGCAGCATATTGGAGGGATCTGTG 160730  
QY 614 GTGCTTGAAGGTTATCAGAGAGGGAGCAGAGTGTGTCTTCCCTTTGTGTAA 673  
Db 160729 GTGCTTGAAGGTTATCAGAGAGGGAGCAGAGTGTGTCTTCCCTTTGTGTAA 160670  
QY 674 CGTAGAGCAGAGGAGAGCTTATGTCAGCAGATTCCTCAAGTGTGGTGGTCCGTTGGTTC 733  
Db 160669 CGTAGAGCAGAGGAGAGCTTATGTCAGCAGATTCCTCAAGTGTGGTGGTCCGTTGGTTC 160610  
QY 734 CCGAGATPACTTTTATGTTGTGTGAGGAGCTTGTGTGAGTGTGCACTTGTGAGTGTGCA 793  
Db 160609 CCGAGA-ACCTTTAGTGTGTATGAGGAGCTTGTGTGAGTGTGCACTTGTGAGTGTGCA 160551  
QY 794 GCGATGCGCTTCCCAATTCCTCTCAATCTCTTTTATGAGCAGAGATCTCAGCTGATGCCA 853  
Db 160550 GCGATGCGCTTCCCAATTCCTCTCAATCTCTTTTATGAGCAGAGATCTCAGCTGATGCCA 160491  
QY 854 ACATGTTCCGATGCTGTGAGAGCATGCCGCTCTCTCTGCTTATGAGGAGCAGAGACTT 913  
Db 160490 ACATGTTCCGATGCTGTGAGAGCATGCCGCTCTCTCTGCTTATGAGGAGCAGAGACTT 160431  
QY 914 GCGCTTGAAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973  
Db 160430 GCGCTTGAAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 160371  
QY 974 ATATTTTATGCTACTTTTATGTTTATGAGCAGAGTGTGAGTGTGAGTGTGAGTGTGAG 1033  
Db 160370 ATATTTTATGCTACTTTTATGTTTATGAGCAGAGTGTGAGTGTGAGTGTGAGTGTGAG 160311  
QY 1034 ACCTTTTCAATTAATAGATTAAAGTAAAA 1063  
Db 160310 ACCTTTTCAATTAATAGATTAAAGTAAAA 160281

RESULT 6  
AR413227 483 bp DNA linear PAT 18-DEC-2003  
LOCUS AR413227  
DEFINITION Sequence 864 from patent US 6639063.  
ACCESSION AR413227  
VERSION AR413227.1 GI:40168337  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 483)  
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.  
TITLE EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 864 28-OCT-2003;  
FEATURES  
source 1..483  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 39.7%; Score 473; DB 6; Length 483;  
Best Local Similarity 100.0%; Pred. No. 3,4e-278; Indels 0; Gaps 0;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 131 TTCTGTGAGGAGCTTCAAGGAGACCAAGGAGGCTGTGAGAGTGTGAGATCTTGGCCAA 190  
QY 181 TGGGGAATTCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
Db 191 TGGGGAATTCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250  
QY 241 TAGCATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
Db 251 TAGCATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 310  
QY 301 CCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
Db 311 CCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370  
QY 361 CAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
Db 371 CAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 430  
QY 421 GCGGTGAGAGCTTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 473  
Db 431 GCGGTGAGAGCTTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 483

RESULT 7  
BD108780 483 bp DNA linear PAT 18-SEP-2002  
LOCUS BD108780  
DEFINITION EST and encoded human protein.  
ACCESSION BD108780  
VERSION BD108780.1 GI:23203598  
KEYWORDS JP 2002010789-A/857.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 483)  
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 857 15-JAN-2002;  
COMMENT OS Homo sapiens (human)  
PN JP 2002010789-A/857  
PD 15-JAN-2002 JP 2000280989  
PF 07-AUG-2000 JP 60/147499  
PI 05-AUG-1999 US 60/147499  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21  
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC EST and encoded human protein  
FH Key Location/Qualifiers  
FT CDS 171..482.  
FT Location/Qualifiers  
1..483  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 39.7%; Score 473; DB 6; Length 483;  
Best Local Similarity 100.0%; Pred. No. 3,4e-278; Indels 0; Gaps 0;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 121 TTGTGTGACAGCCATCAGGAGACCACAGGCGCTGGAGAGATGTGCGATCTTGGCCAA 180
Db 131 TTGTGTGACAGCCATCAGGAGACCACAGGCGCTGGAGAGATGTGCGATCTTGGCCAA 190
QY 181 TGGGGAATGTGTGAGAGACGACACCCCGAGTGAAGACCACTACCAAGCAGCAGGAG 240
Db 191 TGGGGAATGTGTGAGAGACGACACCCCGAGTGAAGACCACTACCAAGCAGGAG 250
QY 241 TAGCATCTCTGACAGAGCTTCTTCAATAGAGGCGCATGTGTCTCCCGAGGAGGTCTCTG 300
Db 251 TAGCATCTCTGACAGAGCTTCTTCAATAGAGGCGCATGTGTCTCCCGAGGAGGTCTCTG 310
QY 301 CCCCCGACAGAGAGGAGGAGGTCAGAGGTCGAGTGTCTGCTCAATGCTTCAATAGACT 360
Db 311 CCCCCGACAGAGAGGAGGAGGTCAGAGGTCGAGTGTCTGCTCAATGCTTCAATAGACT 370
QY 361 CAACCGGACAGTGTGTGAACATGGGCTTTCCGAGTGGCATCTGGGCAACCATCTGTGA 420
Db 371 CAACCGGACAGTGTGTGAACATGGGCTTTCCGAGTGGCATCTGGGCAACCATCTGTGA 430
QY 421 GCGGCTGACCTCCATCTGCTCTCTTCTTCTGCTCATGATGCTTGTGCTG 473
Db 431 GCGGCTGACCTCCATCTGCTCTCTTCTTCTGCTCATGATGCTTGTGCTG 483

RESULT 8
AX113082 550 bp DNA linear PAT 01-MAY-2001
LOCUS AX113082
DEFINITION Sequence 5 from Patent WO0127276.
ACCESSION AX113082
VERSION AX113082.1 GI:13939514
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Harlocker, S.L., Dillon, D.C. and Xu, J.
Dna sequences from breast tumor and uses thereof
Patent: WO 0127276-A 5 19-APR-2001;
CORIXA CORPORATION (US)

FEATURES
source 1..550
location/Qualifiers
1..550
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..550
/note="n = A,T,C or G"

ORIGIN
Query Match 31.0%; Score 369; DB 6; Length 550;
Best Local Similarity 100.0%; Pred.No.2.4e-214;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GCGGCTGACAGACAGCATCTACTAGCGTGGTCACTCTGTGAACATCACTGACTGC 106
Db 1 GCGGCTGACAGACAGCATCTACTAGCGTGGTCACTCTGTGAACATCACTGACTGC 60

QY 107 AAGCTTCCCTCAATTTCTGTGTGAGCCCATCAAGGAGCCCAAGGCTTGGAGAGATGTG 166
Db 61 AAGCTTCCCTCAATTTCTGTGTGAGCCCATCAAGGAGCCCAAGGCTTGGAGAGATGTG 120

QY 167 CGGATCTTGGCAATGGGAAATGTGTGAGAGACAGACCCCGAGTGAAGACCACTACC 226
Db 121 CGGATCTTGGCAATGGGAAATGTGTGAGAGACAGACCCCGAGTGAAGACCACTACC 180

QY 227 CAGCCACCAAGAGTAGATCTCTGACAGAGCTTCTTCAATAGAGGCGCATGTGTCTCC 286
Db 181 CAGCCACCAAGAGTAGATCTCTGACAGAGCTTCTTCAATAGAGGCGCATGTGTCTCC 240

QY 287 CCAAGGGGCTCTGAGCCCGGACAGAGAGGAGGAGGCTGAGGCTGAGTCTCAGTCC 346
Db 241 CCAAGGGGCTCTGAGCCCGGACAGAGAGGAGGAGGCTGAGGCTGAGTCTCAGTCC 300

```

```

QY 347 CCCTTCAATGACCTCAACCGGAGCTGTGAACATGAGGCTTCCGACATGCGATCTCGC 406
Db 301 CCCTTCAATGACCTCAACCGGAGCTGTGAACATGAGGCTTCCGACATGCGATCTCGC 360
QY 407 AACCATGCT 415
Db 361 AACCATGCT 369

RESULT 9
G22561/c 480 bp DNA linear STS 31-MAY-1996
LOCUS G22561
DEFINITION human STS WI-12709, sequence tagged site.
ACCESSION G22561
VERSION G22561.1 GI:1342887
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 480)
Hudson, T.
Whitehead Institute/MIT Center for Genome Research, Physically
Mapped STS
Unpublished (1995)

JOURNAL
COMMENT
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu

Primer A: AAAGGTCACATTAATGATGCC
Primer B: GCTTACGGGAGGTGAAAAA
STS size: 125
PCR Profile:
Presack:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 mM
Tag Polymerase: 0.025 units/ul
Total Vol: 20 ul

Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 9.3

Derived from dbEST (genbank accession H49529).
location/Qualifiers
1..480
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="431.2 CR from top of Chrio linkage group"
25..149
25..149
complement(130..149)

ORIGIN
Query Match 26.1%; Score 311; DB 11; Length 480;
Best Local Similarity 100.0%; Pred.No.9.4e-179;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 713 CCCGAGATCTTTTAAAGTGTATGAGGAGCTGCAATTAAGTGCACAAATGACAGCAAGAA 792  
DB 312 CCCGAGATCTTTTAAAGTGTATGAGGAGCTGCAATTAAGTGCACAAATGACAGCAAGAA 273  
QY 793 AGGATGCGCCCTCCCAATCTCTCAATCTCTTTTATGCGGAGAGATCTGAGTGGATGCC 852  
DB 272 AGGATGCGCCCTCCCAATCTCTCAATCTCTTTTATGCGGAGAGATCTGAGTGGATGCC 213  
QY 853 AACATGTTCCGATCGCTGTGAGACATGCGACGCTCTCTGCTGCTGAGGAGAGAGACT 912  
DB 212 AACATGTTCCGATCGCTGTGAGACATGCGACGCTCTCTGCTGCTGAGGAGAGAGACT 153  
QY 913 TGGGCTTAGGAGAGAGTGAATAAATTCACACTTTTATGACACTGTTTGTGTTTAAATGG 972  
DB 152 TGGGCTTAGGAGAGAGTGAATAAATTCACACTTTTATGACACTGTTTGTGTTTAAATGG 93  
QY 973 TATATTTTATGCTACTTATGTTTATGTTAGACAGAGTGTAGTGGACTTCTATTTATGT 1032  
DB 92 TATATTTTATGCTACTTATGTTTATGTTAGACAGAGTGTAGTGGACTTCTATTTATGT 33  
QY 1033 GACCTTTTCAA 1043  
DB 32 GACCTTTTCAA 22

## RESULT 10

LOCUS AR415427 519 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 3064 from patent US 6639063.  
ACCESSION AR415427  
VERSION AR415427.1 GI:40170537  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 519)  
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.  
TITLE EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 3064 28-OCT-2003;  
FEATURES  
Source location/Qualifiers  
1..519  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 15.5%; Score 185; DB 6; Length 519;  
Best Local Similarity 99.6%; Pred. No. 2.1e-101; Indels 0; Gaps 0;  
Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 58 ACACAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACGCAAGCCTCCCTC 117  
DB 116 ACACAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACGCAAGCCTCCCTC 175  
QY 118 AATTTCTGTGTCAGCCCATCAAGGACCCACAGCGCTCTGGAGAGATGTTGGC 177  
DB 176 AATTTCTGTGTCAGCCCATCAAGGACCCACAGCGCTCTGGAGAGATGTTGGC 235  
QY 178 CAATGGGGAATCGTGAAGACGACGACCCCGAGTGAAGACCACTACCAAGCACAAG 237  
DB 236 CAATGGGGAATCGTGAAGACGACGACCCCGAGTGAAGACCACTACCAAGCACAAG 295  
QY 238 AGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGGCGCATGTGCTCCCGAGGGG 293  
DB 296 AGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGGCGCATGTGCTCCCGAGGGG 351

## RESULT 11

LOCUS BD110980 519 bp DNA linear PAT 18-SEP-2002  
DEFINITION EST and encoded human protein.  
ACCESSION BD110980  
VERSION BD110980.1 GI:23205798

KEYWORDS JP 2002010789-A/3057.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 519)  
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 3057 15-JAN-2002;  
GENSET CORP

## COMMENT

OS Homo sapiens (human)  
PN JP 2002010789-A/3057  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 2000280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC EST and encoded human protein  
CC EST and encoded human protein  
FH Key location/Qualifiers  
FT CDS 93..431.  
location/Qualifiers  
1..519  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## FEATURES

## source

## ORIGIN

Query Match 15.5%; Score 185; DB 6; Length 519;  
Best Local Similarity 99.6%; Pred. No. 2.1e-101; Indels 0; Gaps 0;  
Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 58 ACACAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACGCAAGCCTCCCTC 117  
DB 116 ACACAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACGCAAGCCTCCCTC 175  
QY 118 AATTTCTGTGTCAGCCCATCAAGGACCCACAGCGCTCTGGAGAGATGTTGGC 177  
DB 176 AATTTCTGTGTCAGCCCATCAAGGACCCACAGCGCTCTGGAGAGATGTTGGC 235  
QY 178 CAATGGGGAATCGTGAAGACGACGACCCCGAGTGAAGACCACTACCAAGCACAAG 237  
DB 236 CAATGGGGAATCGTGAAGACGACGACCCCGAGTGAAGACCACTACCAAGCACAAG 295  
QY 238 AGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGGCGCATGTGCTCCCGAGGGG 293  
DB 296 AGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGGCGCATGTGCTCCCGAGGGG 351

## RESULT 12

LOCUS AX431087 643 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 30 from Patent WO0240335.  
ACCESSION AX431087  
VERSION AX431087.1 GI:21656064  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Recipon,H., Liu,C., Salceda,S., Sun,Y., Caffery,R. and Macina,R.A.  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
COMPOSITIONS and methods relating to ovary specific genes and proteins

JOURNAL Patent: WO 0240335-A 30 23-MAY-2002;  
DIADEXUS INC (US)  
FEATURES  
source location/Qualifiers  
1..643  
/organism="Homo sapiens"



|               |                         |
|---------------|-------------------------|
|               | /organelle="mito"       |
|               | /mol_type="genomic DNA" |
|               | /db_xref="taxon:9606"   |
|               | /chromosome="2"         |
|               | /map="2"                |
| repeat_region | /clone="RP11-259N19"    |
|               | /clone_1fb="RPC1-11"    |
|               | /rpt_family="L2"        |
| repeat_region | 280..332                |
|               | /rpt_family="(CA)n"     |
| repeat_region | 1453..1675              |
|               | /rpt_family="L1"        |
| repeat_region | 1755..2031              |
|               | /rpt_family="ERV.L"     |
| repeat_region | 2032..2323              |
|               | /rpt_family="Alu"       |
| repeat_region | 2324..2556              |
|               | /rpt_family="B2RV.L"    |
| repeat_region | 4046..4551              |
|               | /rpt_family="L2"        |
| repeat_region | 4551..4859              |
|               | /rpt_family="MER2_type" |
| repeat_region | 4873..5205              |
|               | /rpt_family="MER2_type" |
| repeat_region | 5212..5251              |
|               | /rpt_family="(CA)n"     |
| repeat_region | 5255..5833              |
|               | /rpt_family="L2"        |
| repeat_region | 6960..7028              |
|               | /rpt_family="L1"        |
| repeat_region | 7029..7327              |
|               | /rpt_family="Alu"       |
| repeat_region | 7328..7393              |
|               | /rpt_family="L1"        |
| repeat_region | 7433..7625              |
|               | /rpt_family="L1"        |
| repeat_region | 7626..7931              |
|               | /rpt_family="Alu"       |
| repeat_region | 7631..7650              |

[illegible]

```
Search completed: August 22, 2004, 06:19:10
Job time : 4693 secs
```

**This Page Blank (uspto)**

OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 03:29:28 ; Search time 514 Seconds

(without alignment)  
9851.852 Million cell updates/sec

Title: US-10-001-885-31

Sequence: 1 aaacgcgcgcgaagcgcgcg.....gaagaggaacacaggaacc 1192

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 212409041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: geneseqn1808.\*
- 2: geneseqn1908.\*
- 3: geneseqn2000.\*
- 4: geneseqn2001.\*
- 5: geneseqn2001b.\*
- 6: geneseqn2002.\*
- 7: geneseqn2003a.\*
- 8: geneseqn2003b.\*
- 9: geneseqn2003c.\*
- 10: geneseqn2004.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 1192  | 100.0       | 1192   | 6  | ABN87820 Human ova  |
| 2          | 903   | 75.8        | 1108   | 6  | ABL90057 Human pol  |
| 3          | 792   | 66.4        | 1049   | 2  | AAV59558 Human sec  |
| 4          | 792   | 66.4        | 1049   | 6  | AB873545 Human CDN  |
| 5          | 792   | 66.4        | 1049   | 8  | ACD82688 CDNA sequ  |
| 6          | 562   | 47.1        | 562    | 6  | ABZ11212 Human pol  |
| 7          | 369   | 31.0        | 550    | 4  | AAE82505 Human bre  |
| 8          | 326   | 27.3        | 365    | 7  | ABX92251 Human ova  |
| 9          | 235   | 19.7        | 517    | 7  | ABX74673 Human CDN  |
| 10         | 195   | 16.4        | 612    | 7  | ABX92250 Human ova  |
| 11         | 161   | 13.5        | 1165   | 5  | AA64820 DNA encod   |
| 12         | 159   | 13.3        | 348    | 5  | AA64819 DNA encod   |
| 13         | 157   | 13.2        | 587    | 5  | AA64821 DNA encod   |
| 14         | 155   | 10.5        | 643    | 6  | ABN87819 Human ova  |
| 15         | 38    | 3.2         | 473    | 4  | AA116223 Probe #61  |
| 16         | 38    | 3.2         | 473    | 4  | AA116277 Probe #61  |
| 17         | 38    | 3.2         | 473    | 4  | ABAS7746 Human foe  |
| 18         | 38    | 3.2         | 473    | 4  | ABAS8927 Human foe  |
| 19         | 38    | 3.2         | 473    | 4  | AA188644 Probe #73  |
| 20         | 38    | 3.2         | 473    | 4  | AA137326 Probe #60  |
| 21         | 38    | 3.2         | 473    | 4  | ABAS77125 Probe #55 |
| 22         | 38    | 3.2         | 473    | 4  | ABAS7795 Probe #62  |
| 23         | 38    | 3.2         | 473    | 4  | AAK32829 Human bon  |

|    |    |     |       |   |                     |
|----|----|-----|-------|---|---------------------|
| 24 | 38 | 3.2 | 473   | 4 | AAK31436 Human bon  |
| 25 | 38 | 3.2 | 473   | 4 | AAK07089 Human bra  |
| 26 | 38 | 3.2 | 473   | 4 | AAK05810 Human bra  |
| 27 | 38 | 3.2 | 473   | 4 | AB832554 Human liv  |
| 28 | 38 | 3.2 | 473   | 4 | AB831118 Human liv  |
| 29 | 38 | 3.2 | 473   | 6 | AB806190 Human gen  |
| 30 | 38 | 3.2 | 473   | 6 | AB807633 Human gen  |
| 31 | 38 | 2.9 | 2836  | 2 | AAV40745 Human gen  |
| 32 | 34 | 2.9 | 2836  | 2 | AAV40744 C. felis   |
| 33 | 34 | 2.9 | 2836  | 2 | AAV40744 C. felis   |
| 34 | 34 | 2.9 | 3288  | 4 | AAAD05613 Human sec |
| 35 | 34 | 2.9 | 3393  | 4 | AAAD05582 Human sec |
| 36 | 33 | 2.8 | 653   | 6 | ABO57023 Human col  |
| 37 | 33 | 2.8 | 1021  | 6 | ABK86007 Rainbow t  |
| 38 | 33 | 2.8 | 1256  | 6 | AAK84605 Human ZFS  |
| 39 | 33 | 2.8 | 43064 | 7 | ACC80906 Human CYP  |
| 40 | 33 | 2.8 | 96595 | 8 | ADA02936 Human BRA  |
| 41 | 33 | 2.8 | 96595 | 9 | ADB72674 Human BRA  |
| 42 | 33 | 2.8 | 96595 | 9 | ADC85416 Human Bra  |
| 43 | 32 | 2.7 | 152   | 4 | AA60109 Human can   |
| 44 | 32 | 2.7 | 157   | 6 | AB186595 Human ova  |
| 45 | 32 | 2.7 | 473   | 6 | ABV95508 Human pan  |

ALIGNMENTS

RESULT 1  
ABN87820  
ID ABN87820 standard; cDNA; 1192 BP.  
XX  
AC ABN87820;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE Human ovary specific nucleic acid SEQ ID NO:31.  
XX  
KW Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;  
KW ovary specific gene; OSO; ovarian cancer; immune response; metastasis;  
KW chromosome 10; gene; se.  
XX  
OS Homo sapiens.  
XX  
PN WO200240535-A2.  
XX  
PD 23-MAY-2002.  
XX  
PF 20-NOV-2001; 2001WO-US045011.  
XX  
PR 20-NOV-2000; 2000US-0252061P.  
XX  
PR 27-NOV-2000; 2000US-0253257P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Salceda S, Macina RA, Recipon H, Cafierkey R, Sun Y, Liu C;  
XX  
XX WPI; 2002-471617/50.  
XX  
DR New ovary specific genes and proteins, useful as a vaccine for treating  
XX patients with ovarian cancer, or for diagnosing and monitoring the  
XX presence and metastases of ovarian cancer in a patient.  
XX  
PT Claim 1; Page 173-174; 260pp; English.  
XX  
PS ABN87790 to ABN87882 represent human ovary specific nucleic acid (OSNA)  
XX sequences, and ABN79297 to ABN79370 represent ovary specific protein  
XX (OSP) sequences from the present invention. OSNA and OSP sequences have  
XX cytostatic activity, and can be used in vaccine production and gene  
XX therapy. An antibody that specifically binds to an OSP can be used for  
XX treating a patient with ovarian cancer, particularly for inducing an  
XX immune response against the ovarian cancer cell expressing the OSNA or  
XX OSP. The OSNA and OSPs can also be used for diagnosing and monitoring  
XX the presence and metastases of ovarian cancer in a patient

XX Sequence 1192 BP; 262 A; 298 C; 349 G; 283 T; 0 U; 0 Other;

Query Match 100.0%; Score 1192; DB 6; Length 1192;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAACCGCTGAGAGCCCGGAGAGTGAAGCGCCGAGAGCCGAGAGTGGCGCTTGACAGACA 60
Db 1 AAACCGCTGAGAGCCCGGAGAGTGAAGCGCCGAGAGCCGAGAGTGGCGCTTGACAGACA 60
QY 61 CAGCATCTACTCAGCGGTGGGTCACTTGTGAAATCACTGACTGCAAGCTTCTCTCAT 120
Db 61 CAGCATCTACTCAGCGGTGGGTCACTTGTGAAATCACTGACTGCAAGCTTCTCTCAT 120
QY 121 TTCTGTGTGAGCCCATCAGAGGACCCACAGAGCCCTGGAGAGATGCTGCGATCTTGSCCAA 180
Db 121 TTCTGTGTGAGCCCATCAGAGGACCCACAGAGCCCTGGAGAGATGCTGCGATCTTGSCCAA 180
QY 181 TGGGGAATGTGTGAGAGCAGACACCCCGAGTGAAGACCACTACCCAGCCACCAAGAG 240
Db 181 TGGGGAATGTGTGAGAGCAGACACCCCGAGTGAAGACCACTACCCAGCCACCAAGAG 240
QY 241 TAGATATTCCTGACAGAGCTTCTTCAATAGGGGCAATGATGCTCCCGCAGGGGGTCTTG 300
Db 241 TAGATATTCCTGACAGAGCTTCTTCAATAGGGGCAATGATGCTCCCGCAGGGGGTCTTG 300
QY 301 CCCCCGACAGAGCAGAGCTTCCAGAGCTGAGTGGTCTGCTCACTCCCTTCAATAGACT 360
Db 301 CCCCCGACAGAGCAGAGCTTCCAGAGCTGAGTGGTCTGCTCACTCCCTTCAATAGACT 360
QY 361 CAACCGGACAGTGTGAACATGAGCTTCCGAGTGGCACTTGGGCAACAGCTGTGGA 420
Db 361 CAACCGGACAGTGTGAACATGAGCTTCCGAGTGGCACTTGGGCAACAGCTGTGGA 420
QY 421 GCCGGTACCTCCATCTGCT 480
Db 421 GCCGGTACCTCCATCTGCT 480
QY 481 CCTGTGTGCTGTCT 540
Db 481 CCTGTGTGCTGTCT 540
QY 541 TAGGGGAGGGGTTGTGAGAGGACCTTGTGGGCTTGTGTGAGAGCAGCATATTTTG 600
Db 541 TAGGGGAGGGGTTGTGAGAGGACCTTGTGGGCTTGTGTGAGAGCAGCATATTTTG 600
QY 601 AGGGGATCTGTGTGTCTTGAAGTATGATCAGAGGGGACCAAGGTGTGTCTTCC 660
Db 601 AGGGGATCTGTGTGTCTTGAAGTATGATCAGAGGGGACCAAGGTGTGTCTTCC 660
QY 661 AGGGGATCTGTGTGTCTTGAAGTATGATCAGAGGGGACCAAGGTGTGTCTTCC 660
Db 661 AGGGGATCTGTGTGTCTTGAAGTATGATCAGAGGGGACCAAGGTGTGTCTTCC 660
QY 721 GATCGTGTGTCTCCGAGATCTTGTAGTGTATGAGGGGCTTGCATTAAGTGGACAAAA 780
Db 721 GATCGTGTGTCTCCGAGATCTTGTAGTGTATGAGGGGCTTGCATTAAGTGGACAAAA 780
QY 781 TGAAGAGCAAGAAAGGATGCTTCCATTTCTCTCAATCTTTTATGCGGAGAAATCT 840
Db 781 TGAAGAGCAAGAAAGGATGCTTCCATTTCTCTCAATCTTTTATGCGGAGAAATCT 840
QY 841 CAGCTGTGATGCAATCTTCCATGCTCTGTGAAAGCAAGCCGAGCTCTCTCTCTCT 900
Db 841 CAGCTGTGATGCAATCTTCCATGCTCTGTGAAAGCAAGCCGAGCTCTCTCTCTCT 900
QY 901 GGGAGCAGAGCTTGGGCTTGAAGGAGTGAAGAAATTCAGACTTTTATGAGCTGTT 960
Db 901 GGGAGCAGAGCTTGGGCTTGAAGGAGTGAAGAAATTCAGACTTTTATGAGCTGTT 960
QY 961 TTGTTTAAGGTAATTTTATTTATGCTTATTTTATTTAGAGCAAGTGTGTGAGAT 1020
Db 961 TTGTTTAAGGTAATTTTATTTATGCTTATTTTATTTAGAGCAAGTGTGTGAGAT 1020
```

```

Db 961 TTGTTTAAGGTAATTTTATTTATGCTTATTTTATTTAGAGCAAGTGTGTGAGAT 1020
QY 1021 TCTATTTATGTGACCTTTTCAATTAATAGATTTTAAATTAATTAATTAATTAATTAAT 1080
Db 1021 TCTATTTATGTGACCTTTTCAATTAATAGATTTTAAATTAATTAATTAATTAATTAAT 1080
QY 1081 AAAAAGCTGTGGGGTTTACCAGGGGCGCAAGGGGGGCGCCCGGGGGGGAATGTGTTTCC 1140
Db 1081 AAAAAGCTGTGGGGTTTACCAGGGGCGCAAGGGGGGCGCCCGGGGGGGAATGTGTTTCC 1140
QY 1141 CGCCCAATTCCTCCCAATTTTGTGACAAATGAAAGAGACACAGGAAACC 1192
Db 1141 CGCCCAATTCCTCCCAATTTTGTGACAAATGAAAGAGACACAGGAAACC 1192
```

## RESULT 2

ABP0057 standard; cDNA; 1108 BP.

ABL90057;

24-MAY-2002 (first entry)

Human polymuonleotide SEQ ID NO 619.

Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor; antineoplastic; anticonvulsant; antibacterial; antifungal; antiparasitic; cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder; neurological disease; infection; human; secreted protein; gene; se.

Homo sapiens.

WO200190304-A2.

29-NOV-2001.

18-MAY-2001; 2001WO-US016450.

19-MAY-2000; 2000US-020515P.

(HUMA-) HUMAN GENOME SCI INC.

Birze CE, Rosen CA;

MP1; 2002-122018/16.

P-PSDB; ABB89648.

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders.

Claim 4; SEQ ID NO 619; 2081bp + Sequence Listing; English.

The invention relates to novel genes (ABL9449-ABL90853) and proteins (ABL89040-ABL90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's disease, anemia, autoimmune thyroiditis, diabetes mellitus, ulcerative colitis; (c) multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (d) wound healing; (e) neurological disorders e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp:wipo.int/pub/published\\_pct\\_sequences](ftp:wipo.int/pub/published_pct_sequences)

Sequence 1108 BP; 234 A; 274 C; 324 G; 271 T; 0 U; 5 Other;  
Query Match 75.8%; Score 903; DB 6; Length 1108;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1053; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 22 AATGAGAGCGCCGAGAGCCGAGAGTGGCCGCTGCAAGACACAGCATCTAAGCGTGGT 81  
DB 8 AATGAGAGCGCCGAGAGCCGAGAGTGGCCGCTGCAAGACACAGCATCTAAGCGTGGT 67  
QY 82 CACCTCTGTGAACATCACTGACTGCAAGCTCCCTCAATTTTGTGTGAGGCCATCAGG 141  
DB 68 CACCTCTGTGAACATCACTGACTGCAAGCTCCCTCAATTTTGTGTGAGGCCATCAGG 127  
QY 142 ACCCAAGGCGCCGAGAGAGTGGTGGGATCTTGGCCAAATGAGGAGAGCA 201  
DB 128 ACCCAAGGCGCCGAGAGAGTGGTGGGATCTTGGCCAAATGAGGAGAGCA 187  
QY 202 CGAAGCCGAGAGAGCACTAACCAGCAACAGAGTACATCTCTGCAAGAGCTT 261  
DB 188 CGAAGCCGAGAGAGCACTAACCAGCAACAGAGTACATCTCTGCAAGAGCTT 247  
QY 262 CTTCAATAGGAGGAGTGGTGGTGGGAGGAGTGGGAGGAGTGGGAGGAGGAGG 321  
DB 248 CTTCAATAGGAGGAGTGGTGGTGGGAGGAGTGGGAGGAGTGGGAGGAGGAGG 307  
QY 322 TGGCAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 381  
DB 308 TGGCAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 367  
QY 382 GGGCTTTCCGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 441  
DB 368 GGGCTTTCCGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 427  
QY 442 CCTCTTCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 501  
DB 428 CCTCTTCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 487  
QY 502 GGTGTCCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561  
DB 488 GGTGTCCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 547  
QY 562 GAGCTTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 621  
DB 548 GAGCTTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 607  
QY 622 AAGGTATGATCAGAGAGGAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681  
DB 608 AAGGTATGATCAGAGAGGAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 667  
QY 682 AAGAGGAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 741  
DB 668 AAGAGGAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 727  
QY 742 CTTTATAGGTATGGGAGCTGATTAAGTGGCAAAATCAGAGCAAGAGAGATGCC 801  
DB 728 CTTTATAGGTATGGGAGCTGATTAAGTGGCAAAATCAGAGCAAGAGAGATGCC 787  
QY 802 CTTGCCAATTCCTCAATCCTTTTATGCGGAGAAATCTGAGTGCATCAATCTTC 861  
DB 788 CTTGCCAATTCCTCAATCCTTTTATGCGGAGAAATCTGAGTGCATCAATCTTC 847  
QY 862 CGAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921  
DB 848 CGAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 907  
QY 922 GGCAGGTGAGAAAAATTCAGAGCTTTTATGAGCACTGTTTGTATATGATATTTT 981  
DB 908 GGCAGGTGAGAAAAATTCAGAGCTTTTATGAGCACTGTTTGTATATGATATTTT 967  
QY 982 ATTGGCTACTTATTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1041  
DB 968 ATTGGCTACTTATTTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027

QY 1042 AATTAATGATTTAAGTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077  
DB 1028 AATTAATGATTTAAGTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1063  
RESULT 3  
AAVS9558  
ID AAVS9558 standard; DNA, 1049 BP.  
XX  
AC AAVS9558;  
XX  
DT 06-JAN-1999 (first entry)  
XX  
DE Human secreted protein gene 48 clone HRCAT74.  
XX  
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
OS Homo sapiens.  
XX  
EN WO9839448-A2.  
XX  
PD 11-SEP-1998.  
XX  
PF 06-MAR-1998; 98WO-US004493.  
XX  
PR 07-MAR-1997; 97US-0038621P.  
XX  
PR 07-MAR-1997; 97US-0040161P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040331P.  
PR 07-MAR-1997; 97US-0040332P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040335P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040337P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043577P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043579P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043673P.  
PR 11-APR-1997; 97US-0043674P.  
PR 11-APR-1997; 97US-0043675P.  
PR 11-APR-1997; 97US-0043676P.  
PR 11-APR-1997; 97US-0043677P.  
PR 11-APR-1997; 97US-0043678P.  
PR 11-APR-1997; 97US-0043679P.  
PR 11-APR-1997; 97US-0043680P.  
PR 11-APR-1997; 97US-0043681P.  
PR 11-APR-1997; 97US-0043682P.  
PR 11-APR-1997; 97US-0043683P.  
PR 11-APR-1997; 97US-0043684P.  
PR 11-APR-1997; 97US-0043685P.  
PR 11-APR-1997; 97US-0043686P.  
PR 11-APR-1997; 97US-0043687P.  
PR 11-APR-1997; 97US-0043688P.  
PR 11-APR-1997; 97US-0043689P.  
PR 11-APR-1997; 97US-0043690P.  
PR 11-APR-1997; 97US-0043691P.  
PR 11-APR-1997; 97US-0043692P.  
PR 11-APR-1997; 97US-0043693P.



PR 23-MAY-1997; 97US-0047594P.  
 PR 23-MAY-1997; 97US-0047595P.  
 PR 23-MAY-1997; 97US-0047596P.  
 PR 23-MAY-1997; 97US-0047597P.  
 PR 23-MAY-1997; 97US-0047598P.  
 PR 23-MAY-1997; 97US-0047599P.  
 PR 23-MAY-1997; 97US-0047600P.  
 PR 23-MAY-1997; 97US-0047601P.  
 PR 23-MAY-1997; 97US-0047612P.  
 PR 23-MAY-1997; 97US-0047613P.  
 PR 23-MAY-1997; 97US-0047614P.  
 PR 23-MAY-1997; 97US-0047615P.  
 PR 23-MAY-1997; 97US-0047616P.  
 PR 23-MAY-1997; 97US-0047617P.  
 PR 23-MAY-1997; 97US-0047618P.  
 PR 23-MAY-1997; 97US-0047632P.  
 PR 23-MAY-1997; 97US-0047633P.  
 PR 06-JUN-1997; 97US-0048964P.  
 PR 06-JUN-1997; 97US-0048974P.  
 PR 13-JUN-1997; 97US-0049610P.  
 PR 08-JUL-1997; 97US-0051926P.  
 PR 16-JUL-1997; 97US-0052874P.  
 PR 18-AUG-1997; 97US-0055724P.  
 PR 22-AUG-1997; 97US-0056530P.  
 PR 22-AUG-1997; 97US-0056631P.  
 PR 22-AUG-1997; 97US-0056632P.  
 PR 22-AUG-1997; 97US-0056637P.  
 PR 22-AUG-1997; 97US-0056638P.  
 PR 22-AUG-1997; 97US-0056642P.  
 PR 22-AUG-1997; 97US-0056643P.  
 PR 22-AUG-1997; 97US-0056644P.  
 PR 22-AUG-1997; 97US-0056842P.  
 PR 22-AUG-1997; 97US-0056843P.  
 PR 22-AUG-1997; 97US-0056844P.  
 PR 22-AUG-1997; 97US-0056877P.  
 PR 22-AUG-1997; 97US-0056878P.  
 PR 22-AUG-1997; 97US-0056879P.  
 PR 22-AUG-1997; 97US-0056880P.  
 PR 22-AUG-1997; 97US-0056881P.  
 PR 22-AUG-1997; 97US-0056882P.  
 PR 22-AUG-1997; 97US-0056884P.  
 PR 22-AUG-1997; 97US-0056886P.  
 PR 22-AUG-1997; 97US-0056887P.  
 PR 22-AUG-1997; 97US-0056888P.  
 PR 22-AUG-1997; 97US-0056889P.  
 PR 22-AUG-1997; 97US-0056892P.  
 PR 22-AUG-1997; 97US-0056893P.  
 PR 22-AUG-1997; 97US-0056894P.  
 PR 22-AUG-1997; 97US-0056903P.  
 PR 22-AUG-1997; 97US-0056908P.  
 PR 22-AUG-1997; 97US-0056909P.  
 PR 22-AUG-1997; 97US-0056910P.  
 PR 22-AUG-1997; 97US-0056911P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057659P.  
 PR 05-SEP-1997; 97US-0057761P.  
 PR 12-SEP-1997; 97US-0058785P.  
 PR 02-OCT-1997; 97US-0061060P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC,  
 PI Bednarik DP, Endress GA, Yu G, Ni J, Feng P, Young PG, Greene JM,  
 PI Ferris AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA,  
 PI Moore PA, Shi Y, Lathier DW, Li Y, Zeng Z, Kyaw H,  
 XX  
 XX WPI: 1998-506364/43.  
 DR P-PSDB; AAW74778.  
 PT

PT New isolated human genes and the secreted polypeptide(s) they encode -  
 XX  
 XX

PT disorders, immune diseases, inflammation or blood disorders.  
 XX  
 XX Claim 1, Page 280; 721p; English.  
 XX  
 CC This sequence represents a nucleic acid molecule designated Gene 48 from  
 CC the human cDNA clone HRC48A14 (deposited as clone ATCC 97899 and ATCC  
 CC 209045) which encodes a secreted human protein. The gene can be used to  
 CC generate fusion proteins by linking to the gene to a human immunoglobulin  
 CC Fc portion (e.g. AAV59502) for increasing the stability of the fused  
 CC protein as compared to the human protein only. The invention relates to  
 CC 186 novel genes and their fragments (nucleic acid sequences: AAV59511-  
 CC V59812; amino acid sequences AAW74731-W75026) which are useful for  
 CC preventing, treating or ameliorating medical conditions e.g. by protein  
 CC or gene therapy. Also, pathological conditions can be diagnosed by  
 CC determining the amount of the new polypeptides in a sample or by  
 CC determining the presence of mutations in the new polynucleotides.  
 CC Specific uses are described for each of the 186 polynucleotides, based on  
 CC which tissues they are most highly expressed in (see AAV59511 for  
 CC described uses)  
 XX

SQ Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

Query Match 66.4%; Score 792; DB 2; Length 1049;  
 Best Local Similarity 99.8%; Pred. No. 5.1e-302;  
 Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 46 TGGCCCTGCGACACAGCATCTACTAGCGTGGGTCACTCTGTGAACATCACTGACTG 105  
 DB 1 TCGCGCTGCGACACAGCATCTACTAGCGTGGGTCACTCTGTGAACATCACTGACTG 60  
 QY 106 CAAGCCTCCCTCAATTGTGTGAGCCCATCAGGACCCACAGCGCTGGAGATGAT 165  
 DB 61 CAAGCCTCCCTCAATTGTGTGAGCCCATCAGGACCCACAGCGCTGGAGATGAT 120  
 QY 166 GGGGATCTTGGCCCAATGGGGAAATGCTGACAGACGACGCCCGGATGAGACCACTAC 225  
 DB 121 GGGGATCTTGGCCCAATGGGGAAATGCTGACAGACGACGCCCGGATGAGACCACTAC 180  
 QY 226 CGAGCACCAGAGGTGACATCTCTGACAGAGCTTCTTAATAGGGCCATGGTGTCC 285  
 DB 181 CGAGCACCAGAGGTGACATCTCTGACAGAGCTTCTTAATAGGGCCATGGTGTCC 240  
 QY 286 CCGAGGGGGTCTTGGCCCGCCGACAGACGACGAGTCCAGGCTGGGTCTGTCACTC 345  
 DB 241 CCGAGGGGGTCTTGGCCCGCCGACAGACGAGTCCAGGCTGGGTCTGTCACTC 300  
 QY 346 CCCCTTCAATGACTCAACCGGCGACGTGGGAAATAGTGGCTTCCGAGTGGCATCTCGG 405  
 DB 301 CCCCTTCAATGACTCAACCGGCGACGTGGGAAATAGTGGCTTCCGAGTGGCATCTCGG 360  
 QY 406 CAACCATGCTGTGAGCGCGTGAACCTCATCTGCTCTTCTCTGCTCATGATGCTTGG 465  
 DB 361 CAACCATGCTGTGAGCGCGTGAACCTCATCTGCTCTTCTCTGCTCATGATGCTTGG 420  
 QY 466 TGTTCGTGGCTCTCTCTGTGTGGCTTGTCTTACCTGTGTGTCCACTGATCAGCGGTG 525  
 DB 421 TGTTCGTGGCTCTCTCTGTGTGGCTTGTCTTACCTGTGTGTCCACTGATCAGCGGTG 480  
 QY 526 ACCTCTGAGGGCTGATGAGGGGTGGGTTTGTGAGAGGGACTTGTGGGGCTTGTGTGAG 585  
 DB 481 ACCTCTGAGGGCTGATGAGGGGTGGGTTTGTGAGAGGGACTTGTGGGGCTTGTGTGAG 540  
 QY 586 AGCAGGCAATATTGAGGGGATCTGTGTGCTTGAAGGTATGATCAGAGGGGACCA 645  
 DB 541 AGCAGGCAATATTGAGGGGATCTGTGTGCTTGAAGGTATGATCAGAGGGGACCA 600  
 QY 646 CAGGTGTGTGTTCCCTTGTGTGTAAGGTGAGGACGAGGAGAGAGTATGAT-CCAGCAT 704  
 DB 601 CAGGTGTGTGTTCCCTTGTGTGTAAGGTGAGGACGAGGAGAGAGTATGAT-CCAGCAT 660  
 QY 705 TT-CCAAAGTGTGGGTGGTCCGTGTGTTCCGAGATACCTTTAGGTGTATGGGGCTTG 763

[illegible]

PR 22-AUG-1997; 97US-0056910P.  
 PR 22-SEP-1997; 97US-0056911P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057659P.  
 PR 05-SEP-1997; 97US-0057761P.  
 PR 12-SEP-1997; 97US-0058785P.  
 PR 02-OCT-1997; 97US-0061060P.  
 PR 06-MAR-1998; 98WO-US004493.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC,  
 PI Bednarik DR, Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM,  
 PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA,  
 PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
 XX WPI, 2002-634796/68.  
 DR P-PSDB; ABG95227.  
 XX  
 PT New isolated human secreted protein for diagnosing, preventing, treating  
 PT or ameliorating medical conditions and used as a food additive or  
 PT preservative.  
 PS  
 PS Example 1; SEQ ID NO 58; 129pp; English.  
 XX  
 CC The invention relates to an isolated protein that is one of 186 human  
 CC secreted proteins, given in the specification, encoded by one of 309 cDNA  
 CC sequences also given in the specification. The protein is used in a  
 CC pharmaceutical composition used to prevent, treat or ameliorate a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to  
 CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
 CC components. The present sequence represents a cDNA derived from a gene  
 CC encoding one of the novel human secreted proteins of the invention. Note:  
 CC This sequence did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docid=6420526b1  
 XX  
 XX Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other:  
 Query Match 66.4%; Score 792; DB 6; Length 1049;  
 Best Local Similarity 99.8%; Pred. No. 5.1e-302;  
 Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Db 241 CCAGAGGGGCTCTGGACCCCGCCAGACAGAGGAGTGTCCAGCTGGATGCTCAGTC 300  
 QY  
 QY 346 CCCCTTCAATGACCTCAACCGGGAGCTGTGTGAATGGGCTTTCCGAGTGGCATCTCGG 405  
 Db 301 CCCCTTCAATGACCTCAACCGGGAGCTGTGTGAATGGGCTTTCCGAGTGGCATCTCGG 360  
 QY 406 CAACCATGCTGTGAGAGCGCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 465  
 Db 361 CAACCATGCTGTGAGAGCGCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
 QY 466 TCTTGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 525  
 Db 421 TCTTGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
 QY 526 ACCCTGAGGCTGATAGGGGTGGGTTGTGTGAGAGGACCTTGCTGGGCTTGTGTGAG 585  
 Db 481 ACCCTGAGGCTGATAGGGGTGGGTTGTGTGAGAGGACCTTGCTGGGCTTGTGTGAG 540  
 QY 586 AGCAGGCAATTTGGAGGGGATCTGTGTGTGCTTTGAAGTATATCAAGAGGGGACCA 645  
 Db 541 AGCAGGCAATTTGGAGGGGATCTGTGTGTGCTTTGAAGTATATCAAGAGGGGACCA 600  
 QY 646 CAGGTGTGTCTTCCCTTTGTGTTAAGCGTGAAGGAGGAGAGCTTGTAGT-CCAGCAT 704  
 Db 601 CAGGTGTGTCTTCCCTTTGTGTTAAGCGTGAAGGAGGAGAGCTTGTAGTCCAGCAT 660  
 QY 705 TT-CCAAAGTGTGGGTGGGTCGTGGTTCGAGATCTTTTGTGTGTGTGTGTGTGTGTGT 763  
 Db 661 TTCCCAAGTGTGGGTGGGTCGTGGTTCGAGATCTTTTGTGTGTGTGTGTGTGTGTGT 720  
 QY 764 CATTAAGTGGCAAAATCAAGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 823  
 Db 721 CATTAAGTGGCAAAATCAAGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
 QY 824 TTATGCCAGAGAAATCTCAGCTGATGCCAATGTCGATGCTGTGAGAGCATGCC 883  
 Db 781 TTATGCCAGAGAAATCTCAGCTGATGCCAATGTCGATGCTGTGAGAGCATGCC 840  
 QY 884 GACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 943  
 Db 841 GACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900  
 QY 944 CTTTCTTGAACCTTTTGTGTTAATGATATTTTATTTATTTATTTATTTATTTATTTAT 1003  
 Db 901 CTTTCTTGAACCTTTTGTGTTAATGATATTTTATTTATTTATTTATTTATTTATTTAT 960  
 QY 1004 ACAAGTGTAGTGGCATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1063  
 Db 961 ACAAGTGTAGTGGCATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1020  
 QY 1064 AAAAAAAAAAAAAA 1077  
 Db 1021 AAAAAAAAAAAAAA 1034  
 RESULT 5  
 ACD82688  
 ID ACD82688 standard; cDNA, 1049 BP.  
 XX  
 XX ACD82688;  
 XX  
 XX 22-SEP-2003 (first entry)  
 DE  
 DE cDNA sequence #48 containing coding region of a human secreted protein.  
 KW Human; secreted protein; hyperproliferative disorder; leukaemia;  
 KW breast cancer; wound; reproductive disorder; blood-related disorder;  
 KW haemophilia; thrombocytopaenia; immunodeficiency; thymic hypoplasia;  
 KW Wiskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;  
 KW graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;  
 KW viral infection; bacterial infection; fungal infection; AIDS; sepsis;  
 KW renal disorder; kidney failure; cardiovascular disorder; cytostatic;  
 KW anemia; necrosis; cerebral ischaemia; congenital heart defect.



Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferris AM,  
 PI Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R,  
 PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
 XX WPI: 2003-521800/49.  
 DR P-PSDB; ABO34421.

XX New genes and its encoded prostate cancer antigen proteins, useful for  
 PT preventing, treating, ameliorating or diagnosing e.g. prostate cancers,  
 PT thymic hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral  
 PT ischemia.

PS Claim 4; SEQ ID NO 58; 260bp; English.

XX The present invention relates to the isolation of novel human secreted  
 CC proteins and the polynucleotide sequences encoding them. The invention  
 CC also discloses vectors, host cells, antibodies, and recombinant methods  
 CC for producing human secreted proteins. The polypeptide and polynucleotide  
 CC sequences for the secreted proteins are useful for preventing, treating,  
 CC ameliorating or diagnosing medical conditions such as hyperproliferative  
 CC disorders (e.g. leukemia or breast cancers), wounds, reproductive  
 CC disorders, blood-related disorders (e.g. haemophilia or  
 CC thrombocytopenia), immunodeficiencies (e.g. Wiskott-Aldrich syndrome or  
 CC thymic hypoplasia), autoimmune disorders (e.g. graft-versus-host disease,  
 CC multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g. asthma),  
 CC viral or bacterial or fungal infections (e.g. AIDS or sepsis), renal  
 CC disorders (e.g. kidney failure), cardiovascular disorders (e.g. angina  
 CC pectoris, cerebral ischaemia or congenital heart defects), respiratory  
 CC disorders, neurological disorders (e.g. Alzheimer's disease or  
 CC Parkinson's disease), and inflammations (e.g. Crohn's disease). The  
 CC polynucleotide or polypeptide may also be used as vaccine adjuvants.  
 CC ACB82641:ACB82950: encode human secreted proteins or their fragments.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO web site at [seqdata.uspto.gov/psipdsidentity.html](http://seqdata.uspto.gov/psipdsidentity.html)

XX Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

Query Match 66.4%; Score 792; DB 8; Length 1049;

Best Local Similarity 99.8%; Pred. No. 5,1e-302;

Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 46 TCGGCGCTGAGACAGACGATCTAATGAGGAGGAGTCACTGTAACATCACTGACTG 105  
 DB 1 TCGGCGCTGAGACAGACGATCTAATGAGGAGGAGTCACTGTAACATCACTGACTG 60  
 QY 106 CAAGCTTCCTCATATTTCTGGTGAAGCCCATCAGGAGCCACAGCGCTGGAGAGATGGT 165  
 DB 61 CAAGCTTCCTCATATTTCTGGTGAAGCCCATCAGGAGCCACAGCGCTGGAGAGATGGT 120  
 QY 166 GCGGATCTGGCCCATGAGGAGAAATCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAG 225  
 DB 121 GCGGATCTGGCCCATGAGGAGAAATCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 QY 226 CCAGCCACCAAGAGAGTGAATTTCTCGACAGAGTTTTCATATAGGAGGAGAGAGAGAG 285  
 DB 181 CCAGCCACCAAGAGAGTGAATTTCTCGACAGAGTTTTCATATAGGAGGAGAGAGAGAG 240  
 QY 286 CCGAGGGGGTCTGGCCCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345  
 DB 241 CCGAGGGGGTCTGGCCCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 QY 346 CCGGTTAATGAGAGTCAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 405  
 DB 301 CCGGTTAATGAGAGTCAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 QY 406 CAACCATGCTGTGAGAGCGGAGTCACTTCCTGCTCTTCTGCTGATGATGCTTGG 465  
 DB 361 CAACCATGCTGTGAGAGCGGAGTCACTTCCTGCTCTTCTGCTGATGATGCTTGG 420  
 QY 466 TGTTCGTGAGGCTCTCTCGTGTGAGGCTGTGATGAGAGAGAGAGAGAGAGAGAGAG 525  
 DB 421 TGTTCGTGAGGCTCTCTCGTGTGAGGCTGTGATGAGAGAGAGAGAGAGAGAGAGAG 480

QY 526 ACCTGTGAGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585  
 DB 481 ACCTGTGAGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 QY 586 AGCAGGATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 645  
 DB 541 AGCAGGATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 QY 646 CAGGAGTGTGTTCCCTGTTGTTAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 704  
 DB 601 CAGGAGTGTGTTCCCTGTTGTTAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 QY 705 TT-CCAAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 763  
 DB 661 TTCCAAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 QY 764 CATTAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 823  
 DB 721 CATTAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 QY 824 TTATGCCGAGAGAGATCTCAGCTGATGCCAATGTTCCGATGCTGTGAGAGAGATGCC 883  
 DB 781 TTATGCCGAGAGAGATCTCAGCTGATGCCAATGTTCCGATGCTGTGAGAGAGATGCC 840  
 QY 884 GACGTTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 943  
 DB 841 GACGTTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
 QY 944 CTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1003  
 DB 901 CTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 QY 1004 ACAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1063  
 DB 961 ACAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
 QY 1064 AAAAAAAAAAAAAA 1077  
 DB 1021 AAAAAAAAAAAAAA 1034

RESULT 6  
 AB211212  
 ID AB211212 standard; cDNA; 562 BP.  
 XX  
 AC AB211212;  
 XX  
 DT 20-JAN-2003 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 94.  
 XX  
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KW arthritis; cytostatic; immunomodulator; neoplastic; neuroprotective;  
 KW antiparasitic; antidiabetic; immunosuppressive; dermatological;  
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
 KW antitumor; gene; ss.  
 KW  
 OS Homo sapiens.  
 XX  
 PN WO200270539-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 05-MAR-2002; 2002WO-US005095.  
 XX  
 PR 05-MAR-2001; 2001US-00799451.  
 XX  
 PA (HYSE-) HYSE INC.  
 XX

PI Tang YT, Zhou P, Goodrich RM, Aarndt V, Zhang J, Zhao Q, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Weinman T, Wang J, Wang D, Drmanac RT;  
XX WPI; 2002-759812/82.  
DR P-PSDB; ABP68995.

XX New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
PT or coagulation disorders.

PS Claim 1; SEQ ID NO 94; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (1) comprising a  
CC nucleotide sequence selected from any of 948 sequences (AB211119-  
CC AB212066) or their mature protein coding portion, active domain coding  
CC protein or complementary sequences. The polynucleotides are useful for  
CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP68902-ABP68949) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell-  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 562 BP; 95 A; 176 C; 174 G; 117 T; 0 U; 0 Other;

Query Match 47.1%; Score 562; DB 6; Length 562;  
Best Local Similarity 100.0%; Pred. No. 1.8e-211;  
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCGGGAGTGAAGCGCCGAGGAGCGGCTGACAGACAGAGTACTGACG 75  
DB 1 GCGGGAGTGAAGCGCCGAGGAGCGGCTGACAGACAGAGTACTGACG 60  
QY 76 GGGGGTCACTCTGTGAATCACTGACAGAGCTCCCTCAATTTGTGGTGGACCCCA 135  
DB 61 GGGGGTCACTCTGTGAATCACTGACAGAGCTCCCTCAATTTGTGGTGGACCCCA 120  
QY 136 TCAGGAGACCAAGGAGCGCTGAGAGATGAGTGGAGATCTTGGCAATGGGAAATGCTGCA 195  
DB 121 TCAGGAGACCAAGGAGCGCTGAGAGATGAGTGGAGATCTTGGCAATGGGAAATGCTGCA 180  
QY 196 GGAAGACGACCCCGAGTGAAGCACTACCCAGCCACCAAGAGTATGATTCCTGACA 255  
DB 181 GGAAGACGACCCCGAGTGAAGCACTACCCAGCCACCAAGAGTATGATTCCTGACA 240  
QY 256 GAGCTTCTTCAATAGAGGAGCAATGAGTCTCCCAAGGAGGCTCTGGCCCGCCAGACAGCA 315  
DB 241 GAGCTTCTTCAATAGAGGAGCAATGAGTCTCCCAAGGAGGCTCTGGCCCGCCAGACAGCA 300  
QY 316 GGAAGTGAAGGAGTGGTGTCTGATGATCCCTTCAATGACCTCAACCGGACAGTGGT 375  
DB 301 GGAAGTGAAGGAGTGGTGTCTGATGATCCCTTCAATGACCTCAACCGGACAGTGGT 360  
QY 376 GAAATGAGGCTTTCCGCAATGAGTCTCCGCAACCAATGAGAGCCGCTGACTCCAT 435  
DB 361 GAAATGAGGCTTTCCGCAATGAGTCTCCGCAACCAATGAGAGCCGCTGACTCCAT 420  
QY 436 CCGGCTCTTCTTCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495  
DB 421 CCGGCTCTTCTTCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 496 CTACCTGCTGTCCCACTGATGACGAGTGAACCTTGAAGGAGTGAAGGAGTGGTGGT 555  
DB 481 CTACCTGCTGTCCCACTGATGACGAGTGAACCTTGAAGGAGTGAAGGAGTGGTGGT 540

QY 556 TGAGAGGAGTCTGCTGGGCTT 577  
DB 541 TGAGAGGAGTCTGCTGGGCTT 562

RESULT 7

ID AAF82505 standard; cDNA; 550 BP.

XX AAF82505;

DT 18-JUN-2001 (first entry)

DE Human breast tumour protein CDNA 15964.

KW Human; breast cancer; tumour; cytostatic; gene therapy; ss.

XX Homo sapiens.

PN WO200127276-A2.

PD 19-APR-2001.

PF 11-OCT-2000; 2000WO-US028255.

PR 12-OCT-1999; 99US-00417031.

PA (CORI-) CORIXA CORP.

PI Harlocker SL, Dillon DC, Xu J;

DR WPI; 2001-273773/28.

PT New polypeptides encoded by polynucleotide sequences over-expressed in  
PT breast tumor tissue are useful to detect, monitor and treat breast  
PT cancer.

PS Claim 13; Page 50; 52pp; English.

CC The present sequence encodes a breast tumour-associated protein. It was  
CC shown to have at least two-fold overexpression in breast tumour tissue.  
CC The invention relates to an isolated polypeptide comprising at least an  
CC immunogenic portion of a breast tumor-specific protein, or its variant  
CC that retains the ability to react with antigen-specific antisera. The  
CC breast tumour polynucleotides, polypeptides and antibodies are useful for  
CC inhibiting development of breast cancer. The polynucleotides may be used  
CC to design primers and probes for detecting and monitoring breast cancer

XX Sequence 550 BP; 102 A; 163 C; 164 G; 118 T; 0 U; 3 Other;

Query Match 31.0%; Score 369; DB 4; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.6e-135;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GCGGCTTGGACACACAGATCTACTGAGGCTGACCTCTGTGAACATCACTGACTGC 106  
DB 1 GCGGCTTGGACACACAGATCTACTGAGGCTGACCTCTGTGAACATCACTGACTGC 60  
QY 107 AAGCTCCCTCAATTTCTGAGAGCCCATCAGGAGCCCAAGGAGGAGGATGATG 166  
DB 61 AAGCTCCCTCAATTTCTGAGAGCCCATCAGGAGCCCAAGGAGGAGGATGATG 120  
QY 167 CGATCTTGGCAATGAGGAAATGTCAGAGACGAGACCCCGAGTGAAGACCACTACC 226  
DB 121 CGATCTTGGCAATGAGGAAATGTCAGAGACGAGACCCCGAGTGAAGACCACTACC 180  
QY 227 CAGGACCAAGAGTATGATCTTCTGACAGAGCTTCTCAATAGGAGGCAATGAGTCTCC 286  
DB 181 CAGGACCAAGAGTATGATCTTCTGACAGAGCTTCTCAATAGGAGGCAATGAGTCTCC 240  
QY 287 CCAAGGAGGCTCTGAGCCCGCCAGACAGAGAGTCCAGAGCTGGTGTCTGCTCAGTCC 346  
DB 241 CCAAGGAGGCTCTGAGCCCGCCAGACAGAGAGTCCAGAGCTGGTGTCTGCTCAGTCC 300

|    |     |                                                            |     |
|----|-----|------------------------------------------------------------|-----|
| QY | 34  | CCCTTCAATGACCTCAACGGGAGCTGGTGAACATAGGGCTTTCCGCAATGCACTCGGC | 406 |
| Db | 301 | CCCTTCAATGACCTCAACGGGAGCTGGTGAACATAGGGCTTTCCGCAATGCACTCGGC | 360 |
| QY | 407 | AACCATGCT                                                  | 415 |
| Db | 361 | AACCATGCT                                                  | 369 |

| RESULT   |                                  |
|----------|----------------------------------|
| 8        | 1                                |
| ABX92251 |                                  |
| ID       | ABX92251 standard; CDNA; 365 BP. |

DT 08-MAY-2003 (first entry)

DE Human ovarian specific nucleic acid DEX0310\_65.

Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer  
non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;  
gynaecological.

**Homo sapiens**

PN W0200292785-A2.

PD 21-NOV-2002.

PF 13-FEB-2002; 2002WO-US022271.

|    |                               |
|----|-------------------------------|
| PR | 13-FEB-2001; 2001US-0268290P. |
| PR | 15-FEB-2001; 2001US-0268834P. |

PA (DIAD-) DIADEXUS INC.

|    |            |            |       |            |          |              |
|----|------------|------------|-------|------------|----------|--------------|
| PI | Salceda S, | Macina RA, | Hu P, | Recipon H, | Karra K, | Cafferkey R, |
| PI | Sun Y,     | Liu C,     |       |            |          |              |

WPI; 2003-120677/11.

PT New isolated OSNA nucleic acid and encoded polypeptide, useful for  
PT identifying, diagnosing, monitoring, staging, imaging and treating  
PT ovarian cancer and non-cancerous diseases in ovarian tissues.

PS Claim 1; Page 185-186; 224pp; English.

CC The invention relates to a new isolated nucleic acid termed ovarian  
CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that  
CC encodes any of 53 fully defined protein sequences appearing as AB61018-  
CC AB61070 (termed ovarian specific proteins), (b) any of 76 fully  
CC defined nucleotide sequences appearing as ABX92187-ABX93262; or (c) a  
CC sequence having at least 60% sequence identity to the nucleic acid  
CC molecule of (a) or (b). Also included are a method for determining the  
CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector  
CC comprising an OSNA, a host cell comprising the vector, an isolated OSP  
CC polypeptide, an anti-OSP antibody or fragment, a method for determining  
CC the presence of an ovary specific protein in a sample and a vaccine  
CC comprising an OSP or OSNA. The methods and compositions of the present  
CC invention are useful for identifying, diagnosing, monitoring, staging,  
CC imaging and treating ovarian cancer and non-cancerous disease in ovary  
CC tissue. The present sequence is an OSNA of the invention

Sequence 365 BP; 60 A; 117 C; 108 G; 80 T; 0 U; 0 Other;

Query Match 27.38; Score 326; DB 7; Length 365;

```
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

200 GACGACCCCGAGTGAGGACCACTACCCAGCCACCAAGAGGTAGCATTCCTCGACAGAGC 259

|    |     |                                                               |     |
|----|-----|---------------------------------------------------------------|-----|
| Db | 40  | GAAGACCCCCGAGTAGAGGACCACTACCCAGGCCACANAGGTATACATTCCTCCAGAGAGC | 99  |
| Qy | 260 | TTCTTTCAATAGAGGGCCATAGTGTCTCCCCAGGGGGTCTGGCCCCGCTCAGACAGAGCA  | 319 |
| Db | 100 | TTCTTCAATAGGGGCCATAGTGTCTCCCCAGGGGGTCTGGCCCCGCTCAGACAGAGCA    | 159 |
| Qy | 320 | GGTGCAGAGCTGGGTCGTCTCAGTCCCCCTTCATATGACCTCAACCCGACAGTGTGAAC   | 379 |
| Db | 160 | GGTGCAGAGCTGGGTCGTCTCAGTCCCCCTTCATATGACCTCAACCCGACAGTGTGAAC   | 219 |
| Qy | 380 | ATGGGCTTTCGGCAGTGGCATCTCCGCAACATGCTGTGAGCCGGGTGACCTCATCTCG    | 439 |
| Db | 220 | ATGGGCTTTCGGCAGTGGCATCTCCGCAACATGCTGTGAGCCGGGTGACCTCATCTCG    | 279 |
| Qy | 440 | CTGCTCTTCTGCTCATGATGCTTGATGTTGATGCTCTGCTCTGCTGTGCTGTATC       | 499 |
| Db | 280 | CTGCTCTTCTGCTCATGATGCTTGATGTTGATGCTCTGCTCTGCTGTGCTGTATC       | 339 |
| Qy | 500 | CTGATGCCCACTAGTCAAGCGGTG                                      | 525 |
| Db | 340 | CTGATGCCCACTAGTCAAGCGGTG                                      | 365 |

RESULT 9  
ABX74673/C

AC ABX74673;

DT 21-MAR-2003 (first entry)

Human cDNA sequence #128 differentially expressed in CC-RCC types.

KW Human; microarray; solid surface; immobilised probe; CC-RCC;

KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;

XX  
OS Homo sapiens.

PN WO200279411-A2

PD 10-OCT-2002.

PF 29-MAR-2002; 2002WO-US009576

PR 29-MAR-2001; 2001US-0279411P

PA (VAND-) VAN ANDEL INST

PI Haab B, Rhodes D, Teh BT, Takashi M;

DR WPI; 2003-040679/03.

PT New microarray, comprising a matrix of cDNA probe from a set of probes  
PT immobilized to a solid surface in predetermined order, useful in the  
PT prognosis of patients with clear cell renal carcinoma.

PS Claim 1; SEQ ID NO 459; 179pp; English.

The present invention relates to a microarray comprising a matrix of at least one cDNA probe from a set of probes immobilised to a solid surface in a predetermined order, where a row of pixels corresponds to replicates of one distinct probe from the set. The probes are complementary to nucleic acid sequences that are expressed differentially in aggressive as compared to non-aggressive types of clear cell renal carcinoma (CC-RCC) and that hybridise to the probes under high stringency conditions. The microarray is useful for the prognosis of patients with CC-RCC, wherein aggressive and non-aggressive CC-RCC tumour types are characterised by differential expression profiles of genes that hybridise with one or more probes immobilised on the microarray. The arrays are useful for gene expression profiling of tumour and normal tissues. The present sequence represents a human cDNA sequence differentially expressed in aggressive



CC vs. non-aggressive type CC-RCC phenotypes  
XX Sequence 517 BP, 166 A, 154 C, 90 G, 107 T, 0 U, 0 Other;  
SQ  
Query Match 19.7%; Score 235; DB 7; Length 517;  
Best Local Similarity 99.7%; Pred. No. 8.8e-83;  
Matches 285; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 503 GTGTCCCACTGAGTCAACCGGTGACCTCTGAGGGGTGATAGCGGTGTTGTTGAGAGG 562  
DB 447 GTGTCCCACTGAGTCAACCGGTGACCTCTGAGGGGTGATAGCGGTGTTGTTGAGAGG 388  
QY 563 GACTTGCTGGGCTTGTGTGAGAGCAGCAGCATATTTGAGGGGATCTGTGTGTGCTTGA 622  
DB 387 GACTTGCTGGGCTTGTGTGAGAGCAGCAGCATATTTGAGGGGATCTGTGTGTGCTTGA 328  
QY 623 AGGTATGATCAGAGAGGGAGCAGCAGGTGTGTGTGTTCCCTTTGTTGTTAAAGCTGAGCA 682  
DB 327 AGGTATGATCAGAGAGGGAGCAGCAGGTGTGTGTGTTCCCTTTGTTGTTAAAGCTGAGCA 268  
QY 683 GAGGAGAGAGTGTAGTCCAGCATTTCCAAAGTGTGGGTGGTCCGTGGTCCGAGATAC 742  
DB 267 GAGGAGAGAGTGTAGTCCAGCATTTCCAAAGTGTGGGTGGTCCGTGGTCCGAGATAC 208  
QY 743 TTTTAGTGTGTATGGGGCTTGCATTAAAGTGCCACAAAATCAGAGCA 788  
DB 207 TTTTAGTGTGTATGGGGCTTGCATTAAAGTGCCACAAAATCAGAGCA 162  
RESULT 10  
ABX92250  
ID ABX92250 standard; cDNA, 612 BP.  
XX  
XX ABX92250;  
XX 08-MAY-2003 (first entry)  
XX  
XX Human ovarian specific nucleic acid DEX0310.64.  
XX  
XX Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;  
XX non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;  
XX gynaecological.  
XX  
XX Homo sapiens.  
XX OS  
XX MO200292785-A2.  
XX PN  
XX  
XX 21-NOV-2002.  
XX PD  
XX  
XX 13-FEB-2002; 2002WO-US022271.  
XX PF  
XX  
XX 13-FEB-2001; 2001US-0268290P.  
XX PR  
XX 15-FEB-2001; 2001US-0268834P.  
XX XX  
XX (DIAD-) DIADEXUS INC.  
XX PA  
XX Salceda S, Macina RA, Hu P, Recipon H, Karra K, Caferkey R;  
XX PI Sun Y, Liu C;  
XX  
XX WPI; 2003-120677/11.  
XX DR P-PSDB; ABU61062.  
XX  
XX  
XX New isolated OSNA nucleic acid and encoded polypeptide, useful for  
XX PT identifying, diagnosing, monitoring, staging, imaging and treating  
XX PT ovarian cancer and non-cancerous diseases in ovarian tissues.  
XX  
XX  
XX Claim 1; Page 185; 224pp; English.  
XX  
XX The invention relates to a new isolated nucleic acid termed ovarian  
XX CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that  
XX CC encodes any of 53 fully defined protein sequences appearing as ABU61018-  
XX CC ABU61070 (termed ovarian specific proteins, OSP); (b) any of 76 fully  
XX CC defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a

CC sequence having at least 60% sequence identity to the nucleic acid  
CC molecule of (a) or (b). Also included are a method for determining the  
CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector  
CC comprising an OSNA, a host cell comprising the vector, an isolated OSP  
CC polypeptide, an anti-OSP antibody or fragment, a method for determining  
CC the presence of an ovary specific protein in a sample and a vaccine  
CC comprising an OSP or OSNA. The methods and compositions of the present  
CC invention are useful for identifying, diagnosing, monitoring, staging,  
CC imaging and treating ovarian cancer and non-cancerous disease in ovary  
CC tissue. The present sequence is an OSNA of the invention  
XX  
XX  
SQ Sequence 612 BP, 156 A, 120 C, 171 G, 165 T, 0 U, 0 Other;  
Query Match 16.4%; Score 195; DB 7; Length 612;  
Best Local Similarity 100.0%; Pred. No. 4.7e-67;  
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 828 GCCGAGAAAGTCTCAGGTGATGCAACATGTTCCAGTGGCTGTGAGAGCAGCCGAGC 887  
DB 261 GCCGAGAAAGTCTCAGGTGATGCAACATGTTCCAGTGGCTGTGAGAGCAGCCGAGC 320  
QY 888 TCTCTCTGCTTGGAGGAGCAGCAGCTTGGGCTTGAAGGAGGTGAAAAAATTCAGACTTT 947  
DB 321 TCTCTCTGCTTGGAGGAGCAGCAGCTTGGGCTTGAAGGAGGTGAAAAAATTCAGACTTT 380  
QY 948 TTTAGCACTGTTTGTGTTTAAATGATATTTTATGCTACTTTATGTTAGAGCA 1007  
DB 381 TTTAGCACTGTTTGTGTTTAAATGATATTTTATGCTACTTTATGTTAGAGCA 440  
QY 1008 GTGGTAGTGCAATTC 1022  
DB 441 GTGGTAGTGCAATTC 455  
RESULT 11  
AAS64820  
ID AAS64820 standard; cDNA, 1165 BP.  
XX  
XX AAS64820;  
XX AC  
XX 13-FEB-2002 (first entry)  
XX DT  
XX  
XX DNA encoding novel human diagnostic protein #624.  
XX DE  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KM food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX XX  
XX Homo sapiens.  
XX OS  
XX MO200175067-A2.  
XX PN  
XX  
XX 11-OCT-2001.  
XX PD  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX PF  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX PR  
XX 23-AUG-2000; 2000US-00649167.  
XX XX  
XX (HYSE-) HYSEQ INC.  
XX PA  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX PI  
XX  
XX WPI; 2001-639362/73.  
XX DR P-PSDB; ABB00633.  
XX  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.  
XX  
XX  
XX Claim 1; SEQ ID NO 624; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX CC

CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1165 BP; 247 A; 279 C; 342 G; 297 T; 0 U; 0 Other;

Query Match 13.5%; Score 161; DB 5; Length 1165;  
Best Local Similarity 100.0%; Pred. No. 1e-53;  
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 903 GACGAGACTGGGCTTAGGGCAGTGGAAAAATCCAGCTTTTTCAGCTGTTT 962  
DB 999 GACGAGACTGGGCTTAGGGCAGTGGAAAAATCCAGCTTTTTCAGCTGTTT 1058  
QY 963 GTTTAATGATATTTTATTGCTACTTTATTGTTTGAACAAGTGTAGTGCATTC 1022  
DB 1059 GTTTAATGATATTTTATTGCTACTTTATTGTTTGAACAAGTGTAGTGCATTC 1118  
QY 1023 TATTTATGACCTTTTCATTAATAGATTAAAGTAAAA 1063  
DB 1119 TATTTATGACCTTTTCATTAATAGATTAAAGTAAAA 1159

RESULT 12

AAS64819/c  
ID AAS64819 standard; cDNA; 348 BP.

AC AAS64819;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #623.

Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG00632.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

PS Claim 1; SEQ ID NO 623; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

SEQ Sequence 348 BP; 121 A; 94 C; 58 G; 75 T; 0 U; 0 Other;

Query Match 13.3%; Score 159; DB 5; Length 348;  
Best Local Similarity 100.0%; Pred. No. 7.6e-53;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 904 AGCAGGACTGGGCTTAGGGCAGTGGAAAAATCCAGCTTTTTCAGCTGTTT 963  
DB 159 AGCAGGACTGGGCTTAGGGCAGTGGAAAAATCCAGCTTTTTCAGCTGTTT 100  
QY 964 TTTTATGATATTTTATTGCTACTTTATTGTTTGAACAAGTGTAGTGCATTC 1023  
DB 99 TTTTATGATATTTTATTGCTACTTTATTGTTTGAACAAGTGTAGTGCATTC 1023  
QY 1024 ATTATTTGACCTTTTCATTAATAGATTAAAGTAAAA 1062  
DB 39 ATTATTTGACCTTTTCATTAATAGATTAAAGTAAAA 1

RESULT 13

AAS64821  
ID AAS64821 standard; cDNA; 587 BP.

AC AAS64821;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #625.

Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

XX





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 04:34:54 ; Search time 2996 Seconds

(without alignments)  
11881.081 Million cell1 updates/sec

Title: US-10-001-885-31

Sequence: 1192  
1 aaacgcctgagccgcgcg.....gagagagacacaggaacc 1192

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc1:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc1:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gse\_hum:\*  
18: em\_gse\_inv:\*  
19: em\_gse\_pin:\*  
20: em\_gse\_vrt:\*  
21: em\_gse\_fun:\*  
22: em\_gse\_man:\*  
23: em\_gse\_mus:\*  
24: em\_gse\_pro:\*  
25: em\_gse\_rtd:\*  
26: em\_gse\_phg:\*  
27: em\_gse\_vr1:\*  
28: gb\_gse1:\*  
29: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| C 1        | 630   | 52.9        | 708    | 12    | BM675290 UI-E-EJO- |
| C 2        | 586   | 49.2        | 688    | 12    | BT521387           |
| C 3        | 510   | 42.8        | 931    | 13    | BM455674           |
| C 4        | 502   | 42.1        | 892    | 13    | BM884216           |

| 5  | 491 | 41.2 | 714  | 10 | BF568253 |
|----|-----|------|------|----|----------|
| 6  | 491 | 41.2 | 1104 | 12 | BG311398 |
| 7  | 484 | 40.6 | 937  | 12 | BM459665 |
| 8  | 482 | 40.4 | 541  | 14 | CA942743 |
| 9  | 480 | 40.3 | 980  | 10 | BE904302 |
| 10 | 475 | 39.8 | 475  | 9  | AI288193 |
| 11 | 464 | 38.9 | 713  | 12 | BI133763 |
| 12 | 455 | 38.2 | 502  | 10 | AM952776 |
| 13 | 446 | 37.4 | 446  | 10 | BG055477 |
| 14 | 443 | 37.2 | 765  | 13 | EX095713 |
| 15 | 431 | 36.2 | 533  | 13 | BU733254 |
| 16 | 427 | 35.8 | 442  | 9  | AA682907 |
| 17 | 424 | 35.6 | 424  | 9  | AA396343 |
| 18 | 423 | 35.5 | 423  | 9  | AA947931 |
| 19 | 423 | 35.5 | 431  | 12 | BM676116 |
| 20 | 423 | 35.5 | 431  | 12 | BM676494 |
| 21 | 423 | 35.5 | 431  | 13 | BU741136 |
| 22 | 422 | 35.4 | 427  | 12 | BM713794 |
| 23 | 416 | 34.9 | 426  | 12 | BM726841 |
| 24 | 411 | 34.5 | 440  | 10 | BE179883 |
| 25 | 410 | 34.4 | 461  | 9  | AA084096 |
| 26 | 401 | 33.6 | 430  | 10 | BF510821 |
| 27 | 395 | 33.1 | 391  | 9  | AI197950 |
| 28 | 391 | 32.8 | 391  | 9  | AI694330 |
| 29 | 390 | 32.7 | 444  | 14 | CB046478 |
| 30 | 386 | 32.4 | 386  | 9  | AI968970 |
| 31 | 376 | 31.5 | 376  | 9  | AI368631 |
| 32 | 374 | 31.4 | 377  | 9  | AM089957 |
| 33 | 369 | 31.0 | 369  | 10 | BF591892 |
| 34 | 361 | 30.3 | 518  | 10 | AM805228 |
| 35 | 354 | 29.7 | 412  | 9  | AI420512 |
| 36 | 352 | 29.5 | 505  | 14 | CA942977 |
| 37 | 338 | 28.4 | 338  | 12 | BM714109 |
| 38 | 337 | 28.3 | 376  | 9  | AA740545 |
| 39 | 321 | 26.9 | 435  | 13 | BU738841 |
| 40 | 311 | 26.1 | 480  | 14 | H49529   |
| 41 | 301 | 25.3 | 425  | 14 | W90083   |
| 42 | 298 | 25.0 | 1117 | 13 | BO891912 |
| 43 | 254 | 21.3 | 528  | 10 | AM960624 |
| 44 | 247 | 20.7 | 510  | 14 | H49767   |
| 45 | 235 | 19.7 | 510  | 9  | AA426015 |

## ALIGNMENTS

RESULT 1  
LOCUS BM675290/C 708 bp mRNA linear EST 27-FEB-2002  
DEFINITION UI-E-EJO-ahr-n-23-0-UI.s1 UI-E-EJO Homo sapiens CDNA clone  
VERSION BM675290  
KEYWORDS BM675290.1 GI:18985188

## ORGANISM

REFERENCE  
AUTHORS Homo sapiens (human)  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 708)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtractions: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)

JOURNAL MEDLINE  
PUBMED  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman

CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 The following repetitive elements were found in this cDNA  
 sequence: 1-42, >AT-richLow\_complexity (matched complement)  
 Seq primer: M13 Forward  
 POLYA=yes.

# FEATURES

source  
 1..708  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-EJ0-ahr-n-23-0-UI"  
 /tissue\_type="fetal eyes, lens, eye anterior segment,  
 optic nerve, retina, Retina Foveal and Macular, RPE and  
 Choroid"  
 /dev\_stage="fetal and adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-EJ0"  
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-EJ0 is a subcloned cDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)<sub>18</sub> tail. The  
 sequence tags for this library are: fetal eyes,  
 AGATCCAGCA; lens, CGATTAGCA; eye anterior segment,  
 AATGCCGCA; optic nerve, CCATTAGTG; retina, CCGCG;  
 Foveal and Macular, GTCC; RPE and Choroid, ACCCA. This  
 library was created for the program, Gene Discovery in the  
 Visual System, supported by National Eye Institute (NEI).  
 TAG TISSUE=Foveal and Macular Retina  
 TAG LIB=UI-E-EJ0  
 TAG\_SEQ=GTCC"

## ORIGIN

Query Match 52.9%; Score 630; DB 12; Length 708;  
 Best Local Similarity 99.9%; Pred. No. 9,7e-118;  
 Matches 680; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

383 GGGTTTCCGAGGCGATCTGCGAACCATGCTGTGAGCCGCTGACCTCCATCTGCTC 442  
 DB GGGTTTCCGAGGCGATCTGCGAACCATGCTGTGAGCCGCTGACCTCCATCTGCTC 642  
 701 GGGTTTCCGAGGCGATCTGCGAACCATGCTGTGAGCCGCTGACCTCCATCTGCTC 642  
 443 CTCTTCCTGCTCAGTATGCTGTGCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTG 502  
 DB CTCTTCCTGCTCAGTATGCTGTGCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTG 582  
 641 CTCTTCCTGCTCAGTATGCTGTGCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTG 582  
 503 GTGTCCCACTGAGTCAAGCGGTGACCTCTGAGGCGGTGATAGGGGTGCTGTTGTAAGG 562  
 DB GTGTCCCACTGAGTCAAGCGGTGACCTCTGAGGCGGTGATAGGGGTGCTGTTGTAAGG 522  
 581 GTGTCCCACTGAGTCAAGCGGTGACCTCTGAGGCGGTGATAGGGGTGCTGTTGTAAGG 522  
 563 GACTTGTGCGCCCTGTGTGAGAGAGAGATTTTGAAGGGGATCTGAGTGGCTTGA 622  
 DB GACTTGTGCGCCCTGTGTGAGAGAGAGATTTTGAAGGGGATCTGAGTGGCTTGA 462  
 521 GACTTGTGCGCCCTGTGTGAGAGAGAGATTTTGAAGGGGATCTGAGTGGCTTGA 462  
 623 AGGATGATCAGAGAGGAGACACAGGTGTGTTCCCTTTGTTAGCGTGAAGCA 682  
 DB AGGATGATCAGAGAGGAGACACAGGTGTGTTCCCTTTGTTAGCGTGAAGCA 402  
 461 AGGATGATCAGAGAGGAGACACAGGTGTGTTCCCTTTGTTAGCGTGAAGCA 402  
 683 GAGGAGACCTTATGCTCAGATTTCCAAAGTGTGGGTGGCTGCTGCTCCGAGATC 742  
 DB GAGGAGACCTTATGCTCAGATTTCCAAAGTGTGGGTGGCTGCTGCTCCGAGATC 342  
 401 GAGGAGACCTTATGCTCAGATTTCCAAAGTGTGGGTGGCTGCTGCTCCGAGATC 342  
 743 TTTTAGGTGATAGGGGCTGCAATTAGTGCACAAATCAGACGAAGCGATGCC 802

DB 341 TTTTAGGTGATAGGGGCTGCAATTAGTGCACAAATCAGACGAAGCGATGCC 282  
 QY 803 TTCCCAATTCCTCAGATCTTTTATGCGAGAGATCTCAGCTGATGCCAATGTTCC 862  
 DB 281 TTCCCAATTCCTCAGATCTTTTATGCGAGAGATCTCAGCTGATGCCAATGTTCC 222  
 QY 863 GATCCCTGTGGAACAACATGCCAGCTCTCTGCGCCATGAGAGAGACCTGGGCTTAG 922  
 DB 221 GATCCCTGTGGAACAACATGCCAGCTCTCTGCGCCATGAGAGAGACCTGGGCTTAG 162  
 QY 923 GCAGGTGGAACAAATTCAGACTTTTATGACACTGTTTGTATGATATATTTTA 982  
 DB 161 GCAGGTGGAACAAATTCAGACTTTTATGACACTGTTTGTATGATATATTTTA 102  
 QY 983 TTGGCTACTTATTTTATGAGACAGGTGATGATGATTTATTTATGACCTTTTCA 1042  
 DB 101 TTGGCTACTTATTTTATGAGACAGGTGATGATGATTTATTTATGACCTTTTCA 42  
 QY 1043 ATTAATGATTTAGTAAAA 1063  
 DB 41 ATTAATGATTTAGTAAAA 21

RESULT 2  
 B1521387  
 LOCUS  
 DEFINITION  
 B1521387 688 bp mRNA linear EST 29-AUG-2001  
 603081177F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5220570 5',  
 mRNA sequence.  
 B1521387  
 B1521387.1 GI:15346179  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 688)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strauberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.lnl.gov>  
 Plate: LAM11554 row: 1 column: 19  
 High quality sequence, stop: 688.

## FEATURES

source  
 1..688  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5220570"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_120"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 PCMV-SPO6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 Kb, insert size range 1-2.5 Kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 49.2%; Score 586; DB 12; Length 688;  
 Best Local Similarity 99.7%; Pred. No. 6,6e-109;  
 Matches 686; Conservative 0; Mismatches 2; Indels 0; Gaps 0;





| SOURCE                                                                                                                      | ORGANISM     | REFERENCE          | TITLE | JOURNAL | COMMENT |
|-----------------------------------------------------------------------------------------------------------------------------|--------------|--------------------|-------|---------|---------|
| Homo sapiens (human)                                                                                                        | Homo sapiens | 1 (bases 1 to 892) |       |         |         |
| Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |              | 1 (bases 1 to 892) |       |         |         |
| NIH-MGC                                                                                                                     |              | 1 (bases 1 to 892) |       |         |         |
| National Institutes of Health, Mammalian Gene Collection (MGC)                                                              |              |                    |       |         |         |
| Unpublished (1999)                                                                                                          |              |                    |       |         |         |
| Contact: Robert Strausberg, Ph.D.                                                                                           |              |                    |       |         |         |

Plate: LLAM13607 row: 1 column: 12  
High quality sequence stop: 497.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source   | 1. .892             |

```

organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGS:6197459"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="VDH108"
/clone_lib="lunpsi_sciatic_nerve"
/note="vector: pGMW-Sport6 (Life Technologies); Site_1:
NcoI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:-
5'-GTCAGCTTCATGAGCGAGCG-3' and
5'-GCTGACCTTCGAGCGCG-3'
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lunpsi, M.D./Ph.D. (Baylors
College of Medicine) and is available through Life
Technologies."

```

ORIGIN

|             |        |            |        |             |
|-------------|--------|------------|--------|-------------|
| Query Match | 42.18; | Score 502; | DB 13; | Length 892; |
|-------------|--------|------------|--------|-------------|

Matches 602; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| Qy | 292 | GGGTCCCTGGACCCCGCAGCAGCAGCAGGTCGCAAGGCTGGGGTCTGCTCAAGTCCCCTT | 351 |
| Db | 1   | GGGTCTGTGCCCCCGCAGCAGCAGGCAAGTCCAGGCTGGGTCTGCTCAAGTCCCCTT    | 60  |
| Qy | 352 | CAATGACTCAACCGCAGCTGTGTAAATGAGGCTTTCCGAGTGGCACTCTCGGCACCA    | 411 |
| Db | 61  | CAATGACTCAACCGCAGCTGTGTAAATGGGCTTTCCGAGTGGCACTTTGGCAACA      | 120 |
| Qy | 412 | TGCTGTGAAGCCGGTGACCTTCATCTGTCCTCTTCTGTCTCAATGATGTTGTGTG      | 471 |
| Db | 121 | TGCTGTGAAGCCGGTGACCTTCATCTCTCTCTCTCTGTGTCAATGATGTTGTGTG      | 180 |
| Qy | 472 | TGGCTCTCTCTCTGTGTGGCTTGTGTCTTACTCTGTGTCCACCTGATGACGGCGTGA    | 531 |
| Db | 181 | TGGCTCTCTCTCTGTGTGGCTTGTGTCTTACTCTGTGTCCACCTGATGACGGCGTGA    | 240 |
| Qy | 532 | GAGGCGTGAATGAGGAGGTGTGTGTGAAGAGGACTTGCTGGGCTTGTGTGTGAGACAG   | 591 |
| Db | 241 | GAGGCGTGAATGAGGAGGTGTGTGTGTGAAGAGGACTTGCTGGGCTTGTGTGTGAGACAG | 300 |
| Qy | 592 | CATATTTGAGGAGGATCTGTGTGTCCTTGAAGGTATGATCAGAGGAGCCACAGGTG     | 651 |
| Db | 301 | CATATTTGAGGAGGATCTGTGTGTCCTTGAAGGTATGATCAGAGGAGCCACAGGTG     | 360 |
| Qy | 652 | TGTGTTCCCTTGTGTGTAAAGCGTAGGCGAGAGGAGACGTTTATGTCACATTTCCAA    | 711 |
| Db | 361 | TGTGTTCCCTTGTGTGTAAAGCGTAGGCGAGAGGAGACGTTTATGTCACATTTCCAA    | 420 |

[illegible]

|            |                                                                  |      |        |                 |
|------------|------------------------------------------------------------------|------|--------|-----------------|
| RESULT 5   | 714 bp                                                           | mRNA | linear | EST 12-DEC-2000 |
| BF568253   |                                                                  |      |        |                 |
| LOCUS      |                                                                  |      |        |                 |
| DEFINITION | 602184625P1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4300518 5', |      |        |                 |
| ACCESSION  | BF568253                                                         |      |        |                 |
| VERSION    | BF568253.1                                                       |      |        |                 |
| KEYWORDS   | EST.                                                             |      |        |                 |
| SOURCE     | Homo sapiens (human)                                             |      |        |                 |
| ORGANISM   | Homo sapiens                                                     |      |        |                 |

|           |                                                                       |
|-----------|-----------------------------------------------------------------------|
| REFERENCE | 1 (bases 1 to 714)                                                    |
| AUTHORS   | NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> |
| TITLE     | National Institutes of Health, Mammalian Gene Collection (MGC).       |
| JOURNAL   | Unpublished (1999)                                                    |
| COMMENT   | Contact: Robert Strausberg, Ph.D.                                     |

Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L10M1159 row: j column: 07  
High quality sequence stop: 712.

## FEATURES

### Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4300518"
/tissue_type="hepithelial carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 42"
/name="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI-drf sites using the
following 5' adaptor: GCCACGAC(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library. |"

```

**ORIGIN**

|                           |       |                   |          |            |
|---------------------------|-------|-------------------|----------|------------|
| Query Match               | 41.2% | Score 491         | DB 10    | Length 714 |
| Best Local Similarity     | 99.8% | Pred. No. 7.2e-90 |          |            |
| Matches 541; Conservative | 0     | Mismatches 1      | Indels 0 | Gaps 0     |

QY 10 GGAGCCCGCCGGGAGTGCACGCCGCCGAGGTGGCGCCTTGCAACAACAGCATCTA 69  
|||  
Db 2 GGAGCCCGCCGGGAGTGCACGCCGCCGAGGTGGCGCCTTGCAACAACAGCATCTA 61



Tissue Procurement: ATCC  
 cDNA library Preparation: Life Technologies, Inc.  
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LIML at:  
<http://image.jnl.gov>  
 Place: LMAN1220 row: 1 column: 16  
 High quality sequence stop: 585.

## FEATURES

source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5534295"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.1 kb."

## ORIGIN

Query Match 40.6%; Score 484; DB 12; Length 937;  
 Best Local Similarity 99.5%; Pred. No. 1.4e-88;  
 Matches 634; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 328 GCTGGGTCCTGCTCACTCCCTCTTCAATCACTCAACCGGACGCTGGTGAACATGGGCTT 387
DB 1 GCTGGGTCCTGCTCACTCCCTCTTCAATCACTCAACCGGACGCTGGTGAACATGGGCTT 60
QY 388 TCCGACAGTGGACATCTCGGACCAACATGCTGGAGCGCGGTGACCTTCATCTGCTCTT 447
DB 61 TCCGACAGTGGACATCTCGGACCAACATGCTGGAGCGCGGTGACCTTCATCTGCTCTT 120
QY 448 CCTGCTCATGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 507
DB 121 CCTGCTCATGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180
QY 508 CCACTGATGACGCGGTGACCTCTGAGGGCTGATAGGGGTGGGTTGTTGAGAGGACTT 567
DB 181 CCACTGATGACGCGGTGACCTCTGAGGGCTGATAGGGGTGGGTTGTTGAGAGGACTT 240
QY 568 GCTGGGCTTGTGTGAGAGCAAGGCAATTTGAGAGGGGATCTGTGTGCTTGAAGTA 627
DB 241 GCTGGGCTTGTGTGAGAGCAAGGCAATTTGAGAGGGGATCTGTGTGCTTGAAGTA 300
QY 628 TGTATGAGAGGGGACCAAGGTGTGTTTCCCTTGTGTTAAGCGTGAAGGAGG 687
DB 301 TGTATGAGAGGGGACCAAGGTGTGTTTCCCTTGTGTTAAGCGTGAAGGAGG 360
QY 688 AGAGCTTAGTCCAGATTTCCAAAGTGTGGTGGGTCCGTTGGTCCCGAGTACTTTTA 747
DB 361 AGAGCTTAGTCCAGATTTCCAAAGTGTGGTGGGTCCGTTGGTCCCGAGTACTTTTA 420
QY 748 GGTGTATGGGGCTCTGATTAAGTGGCAAAATCAGAGCAAGAAAGCAATGCCCTTCCC 807
DB 421 GGTGTATGGGGCTCTGATTAAGTGGCAAAATCAGAGCAAGAAAGCAATGCCCTTCCC 480
QY 808 AATTCCTCAATCTTTTATGCGGAGAAATCTCAGCTGATGCCAATGTTCCGATGC 867
DB 481 AATTCCTCAATCTTTTATGCGGAGAAATCTCAGCTGATGCCAATGTTCCGATGC 540
QY 868 CTGTGAGAGCATGCGAGCTCTCTCTGCTTGAAGGAGCAAGACTTGGGCTTGAAGGAGG 927
DB 541 CTGTGAGAGCATGCGAGCTCTCTCTGCTTGAAGGAGCAAGACTTGGGCTTGAAGGAGG 600
QY 928 TGAATAAATTCAGACTTTTATAGCACTGTTTGT 964
DB 601 TGAATAAATTCAGACTTTTATAGCACTGTTTGT 637

```

CA942743/c  
 LOCUS 541 bp mRNA linear EST 30-DEC-2002  
 DEFINITION i166f05.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6607449 3',  
 mRNA sequence.  
 CA942743  
 CA942743.1 GI:27431223  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 541)  
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
 Schmitt, A., Theising, B., Ritzer, E., Ronko, I., Bennett, J.,  
 Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarelashvili, R.,  
 Williams, T., Jackson, Y., and Bowers, Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)  
 Other ESTs: i166f05.y1  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Harvard University, Howard Hughes Medical Institute  
 Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
 Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: Dr. Hiroshi Inoue  
 (hinoue@im.wustl.edu)  
 Seq primer: -40RP from Glibco  
 High quality sequence stop: 455.  
 Location/Qualifiers  
 1..541  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6607449"  
 /tissue\_type="Purified pancreatic islet"  
 /lab\_host="DH10B"  
 /clone\_1b="HR85 islet"  
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:  
 NotI; Site 2: XhoI; cDNA made by oligo-dT priming.  
 Size-selected on agarose gel. Average insert size -1kb. 5'  
 XhoI site was destroyed after directional cloning.  
 Amplified once. Contact information: Hiroshi Inoue, MD,  
 Metabolism Div. (Alan Permutt Lab), Washington University  
 School of Medicine, Box 8127, 660 South Euclid Ave., St.  
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
 314-362-1916, Fax: 314-747-2692."

## FEATURES

source

Query Match 40.4%; Score 482; DB 14; Length 541;  
 Best Local Similarity 99.8%; Pred. No. 5.9e-88;  
 Matches 532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 531 TGAAGGCTGATAGAGGGTGGGTTGTTGAGAGGACTTGGGCTTGTGTAGAGCGAG 590
DB 541 TGAAGGCTGATAGAGGGTGGGTTGTTGAGAGGACTTGGGCTTGTGTAGAGCGAG 482
QY 591 GCATATTTGAGAGGGGATCTGTGTGCTTGAAGGTATGATCAGAGGGGACCAAGGT 650
DB 481 GCATATTTGAGAGGGGATCTGTGTGCTTGAAGGTATGATCAGAGGGGACCAAGGT 422
QY 651 GTGTGTTCCCTTTGTGTTAAGCGTGAAGGAGGAGGAGCGTTAGTCCAGCATTTCCAA 710
DB 421 GTGTGTTCCCTTTGTGTTAAGCGTGAAGGAGGAGGAGCGTTAGTCCAGCATTTCCAA 362
QY 711 AGTGTGGTGGTCCCTTGTGTTCCAGATCTTTTATAGTGTGATGGGCGCTCATTAAG 770

```

Db 361 AGTGTGGGTGGTCCGTTGGTTCCAGATCTTTTAGGTATGGGGCCCTGCAATTAAG 302  
QY 771 TGGCACAAAATCAAGAGCAAGAAAGATGCCCTTCCCAATTCCTCAATCCTTTATGCG 830  
Db 301 TGGCACAAAATCAAGAGCAAGAAAGATGCCCTTCCCAATTCCTCAATCCTTTATGCG 242  
QY 831 GAGAAAGATCTGAGCTGATGCCCAATTCCTCAATTCCTGGAAGATCCGACGCT 890  
Db 241 GAGAAAGATCTGAGCTGATGCCCAATTCCTCAATTCCTGGAAGATCCGACGCT 182  
QY 891 CCTGTGCTTGGAGAGCAGACCTTGGGCTTAAAGGAGTGAAGAAATTCAGACTTTT 950  
Db 181 CCTGTGCTTGGAGAGCAGACCTTGGGCTTAAAGGAGTGAAGAAATTCAGACTTTT 122  
QY 951 AGCACTGTTTTTTTAAATGATATTTTATTTTATTTGCTACTTTATTTGTTTGGCAAGT 1010  
Db 121 AGCACTGTTTTTTTAAATGATATTTTATTTTATTTGCTACTTTATTTGTTTGGCAAGT 62  
QY 1011 GGAGTGGCAATTCATTTATTTGATCCTTTCAATTAATAGATTAAATAA 1063  
Db 61 GTAGTGGCAATTCATTTATTTGATCCTTTCAATTAATAGATTAAATAA 9

RESULT 9  
BE904302 980 bp mRNA linear EST 20-OCT-2000  
LOCUS 601494571P2 NIH\_MGC\_70 Homo sapiens CDNA clone IMAGE:3896656 5',  
DEFINITION mRNA sequence.  
ACCESSION BE904302  
VERSION BE904302.1 GI:10396415  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 980)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1AM9689 row: n column: 17  
High quality sequence stop: 711.  
Location/Qualifiers  
1..980  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3896656"  
/tissue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: pancreas; Vector: PCMV-SPORE; Site 1: NotI;  
Site 2: SalI. Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."

FEATURES

source  
1..980  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3896656"  
/tissue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: pancreas; Vector: PCMV-SPORE; Site 1: NotI;  
Site 2: SalI. Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."

ORIGIN

Query Match 40.3%; Score 480; DB 10; Length 980;  
Best Local Similarity 99.8%; Pred. No. 8.5e-88;  
Matches 530; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 22 AGTGAAGCCGCGAGGCGCGAGTGGCGCTGAGACACAGATCTTACAGCGTGGT 81  
Db 1 AGTGAAGCCGCGAGGCGCGAGTGGCGCTGAGACACAGATCTTACAGCGTGGT 60

QY 82 CACCTGTGAACATCACTGACGAGCCCTCCCTCAATTTCTGTGACGCCCATCAGG 141  
Db 61 CACCTGTGAACATCACTGACGAGCCCTCCCTCAATTTCTGTGACGCCCATCAGG 120  
QY 142 ACCCAAGGCGCTGGAGATGTGTGGATCTTTGGCAATGGGGAAATGTGAGAGCA 201  
Db 121 ACCCAAGGCGCTGGAGATGTGTGGATCTTTGGCAATGGGGAAATGTGAGAGCA 180  
QY 202 CGACCCCGAGTGAAGACCACTACCCAGCAACCAAGGTAGATCTCTGACAGACT 261  
Db 181 CGACCCCGAGTGAAGACCACTACCCAGCAACCAAGGTAGATCTCTGACAGACT 240  
QY 262 CTTCATAGAGGCGCATGTGTCTCCCAAGGGGCTCTGGCCCCCGACAGAGAGG 321  
Db 241 CTTCATAGAGGCGCATGTGTCTCCCAAGGGGCTCTGGCCCCCGACAGAGAGG 300  
QY 322 TGGCAGGCTGGGTGCTGCTGATCCCCCTTCATGACCTCAACCGAGCTGTGAACAT 381  
Db 301 TGGCAGGCTGGGTGCTGATCCCCCTTCATGACCTCAACCGAGCTGTGAACAT 360  
QY 382 GGGCTTTCGAGTGGCATCTGGCAACCATGTGTGAGACCGGTGACTCTGCT 441  
Db 361 GGGCTTTCGAGTGGCATCTGGCAACCATGTGTGAGACCGGTGACTCTGCT 420  
QY 442 CTTCTTCTGCTCATATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 501  
Db 421 CTTCTTCTGCTCATATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480  
QY 502 GGTGTCCACCTGAGTCAAGCGGTGACCTTGAGGGCTGATAGGGGTGTT 552  
Db 481 GGTGTCCACCTGAGTCAAGCGGTGACCTTGAGGGCTGATAGGGGTGTT 531

RESULT 10  
A1288193 475 bp mRNA linear EST 29-JAN-1999  
LOCUS c190e03.x1 Soares NHHMPu S1 Homo sapiens CDNA clone IMAGE:1879612  
DEFINITION 3' similar to contains MER22. t1 MER22 repetitive element ;, mRNA  
sequence.  
ACCESSION A1288193  
VERSION A1288193  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 475)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 1129 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 409.  
Location/Qualifiers  
1..475  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1879612"  
/tissue\_type="pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/note="Organ: mixed (see below); Vector: pT7T3D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI. Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NbhV, pregnant uterus  
NDHPU, and fetal heart NDH19W) were mixed, and ss circles

FEATURES

source  
1..475  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1879612"  
/tissue\_type="pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/note="Organ: mixed (see below); Vector: pT7T3D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI. Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NbhV, pregnant uterus  
NDHPU, and fetal heart NDH19W) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

## ORIGIN

Query Match 39.8%; Score 475; DB 9; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.7e-86;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 588 CAGCATATTTGGAGGGGATCTGTGTGCTTGAAGATATGATCAGAGGGGACCA 647  
DB 475 CAGCATATTTGGAGGGGATCTGTGTGCTTGAAGATATGATCAGAGGGGACCA 416  
QY 648 GGTGTGTGTTCCCTTGTGTAGGGTGAAGGAGAGCGTATGCCAGATTTC 707  
DB 415 GGTGTGTGTTCCCTTGTGTAGGGTGAAGGAGAGCGTATGCCAGATTTC 356  
QY 708 CAAAGTGTGGTGGGTCGTTGGTCCGAGATCTTTAGGTGATGGGGCTGCATT 767  
DB 355 CAAAGTGTGGTGGGTCGTTGGTCCGAGATCTTTAGGTGATGGGGCTGCATT 296  
QY 768 AAGTGGCAAAATCAGAGCAAGAGCAATGCCCTTCCCAATCTCTCAATCTTTAT 827  
DB 295 AAGTGGCAAAATCAGAGCAAGAGCAATGCCCTTCCCAATCTCTCAATCTTTAT 236  
QY 828 GCCAGAGATCTCAGTGTGATCCAAATGTTCCGATCCGTGGAGACATGCCAG 887  
DB 235 GCCAGAGATCTCAGTGTGATCCAAATGTTCCGATCCGTGGAGACATGCCAG 176  
QY 888 TCTCCTCTGCTTGGAGAGCAGACTGGGCTTAGGGCAGGTGAAAAAATCCAGACTTT 947  
DB 175 TCTCCTCTGCTTGGAGAGCAGACTGGGCTTAGGGCAGGTGAAAAAATCCAGACTTT 116  
QY 948 TTAGACATGTTTTTTTAAATGATATTTTATTTGCTACTTATGTTAGACA 1007  
DB 115 TTAGACATGTTTTTTTAAATGATATTTTATTTGCTACTTATGTTAGACA 56  
QY 1008 GTGTGTGTGCTATTTATTTGATCTTTCAATAAATAGATTAAAGTAAA 1062  
DB 55 GTGTGTGTGCTATTTATTTATTTGATCTTTCAATAAATAGATTAAAGTAAA 1

RESULT 11  
BI193763 713 bp mRNA linear EST 10-JUL-2001  
LOCUS 602948418F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:5091412 5',  
DEFINITION mRNA sequence.  
ACCESSION BI193763  
VERSION BI193763  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 713)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L10M1855 row: d column: 05  
High quality sequence stop:709

## source

1.. 713  
/organism="Homo sapiens"  
/db\_type="mRNA"  
/mol\_type="cDNA"  
/clone="IMAGE:5091412"  
/tissue\_type="epithelioid carcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_42"  
/note="Organ: pancreas; Vector: pOT7; Site: 1: XhoI;  
Site: 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

Query Match 38.9%; Score 464; DB 12; Length 713;  
Best Local Similarity 99.8%; Pred. No. 1.9e-84;  
Matches 514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 142 ACCCAGAGGCTCTGGAGATGTCGGATCTTGGCAATGGGAAATCTGACAGCA 201  
DB 12 ACCCAGAGGCTCTGGAGATGTCGGATCTTGGCAATGGGAAATCTGACAGCA 71  
QY 202 CGACCCCGAGTGAAGCACTACCCAGCAAGAGATGAGATTCTGACAGAGCTT 261  
DB 72 CGACCCCGAGTGAAGCACTACCCAGCAAGAGATGAGATTCTGACAGAGCTT 131  
QY 262 CTTCAATGAGGAGGATGCTCTCCCGAGGGGTCCTGAGCCCGCAGAGAGGCA 321  
DB 132 CTTCAATGAGGAGGATGCTCTCCCGAGGGGTCCTGAGCCCGCAGAGAGGCA 191  
QY 322 TGCCAGGCTGGTGTCTAGTCCCTTCAATGACCTAACCCGAGATGTGAACAT 381  
DB 192 TGCCAGGCTGGTGTCTAGTCCCTTCAATGACCTAACCCGAGATGTGAACAT 251  
QY 382 GGGCTTTCCGAGTGGATCTGCAACCATGCTGTGAGCCGAGTCACTCCCTGCT 441  
DB 252 GGGCTTTCCGAGTGGATCTGCAACCATGCTGTGAGCCGAGTCACTCCCTGCT 311  
QY 442 CTTCTTCTCTCATATGATGTTGTGTGCTGAGCTCTCTGCTGCTGCTGCTACT 501  
DB 312 CTTCTTCTCTCATATGATGTTGTGTGCTGAGCTCTCTGCTGCTGCTGCTACT 371  
QY 502 GGTGTCCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 561  
DB 372 GGTGTCCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 431  
QY 562 GACTTGTCTGGGCTTGTGTGTGAGAGAGGATATTTGAGAGGATCTGAGTGT 621  
DB 432 GACTTGTCTGGGCTTGTGTGTGAGAGAGGATATTTGAGAGGATCTGAGTGT 491  
QY 622 AAGTATGATCAGAGAGGAGACCAAGTGTGTGT 656  
DB 492 AAGTATGATCAGAGAGGAGACCAAGTGTGTGT 526

RESULT 12  
AM952776 502 bp mRNA linear EST 01-JUN-2000  
LOCUS AM952776  
DEFINITION EST164846 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.  
ACCESSION AM952776  
VERSION AM952776.1 GI:8142459  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 502)

TITLE  
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.  
JOURNAL  
Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray  
COMMENT  
Unpublished (2000)  
Contact: John Quackenbush  
The Institute for Genomic Research  
7112 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johng@tigr.org  
Plate: 43  
Seq primer: Reverse.

FEATURES  
source  
Location/Qualifiers  
1..502  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE sequences, YAGB"  
/note="Vector: pBluescriptSKm"

#### ORIGIN

Query Match 38.2%; Score 455; DB 10; Length 502;  
Best Local Similarity 100.0%; Pred. No. 1.7e-82;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 378 ACATGGGCTTTCGGCAGTGGCATCTGGCAACCATCTGTGAGCCGGTGAACCTTCATCC 437  
DB 1 ACATGGGCTTTCGGCAGTGGCATCTGGCAACCATCTGTGAGCCGGTGAACCTTCATCC 60  
QY 438 TGGCTCTCTTCTCGCTCATGATGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 497  
DB 61 TGGCTCTCTTCTCGCTCATGATGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120  
QY 498 ACTGTGTGCCACCTGAGTCAAGCGGTGACCTCTGAGGCTGATAGGGGTGGCTTTGTTG 557  
DB 121 ACTGTGTGCCACCTGAGTCAAGCGGTGACCTCTGAGGCTGATAGGGGTGGCTTTGTTG 180  
QY 558 AGAGGAGCTTGTGGGCTTGTGTGAGAGAGAGCATATTGAGAGGAGCATCTGTGTGTC 617  
DB 181 AGAGGAGCTTGTGGGCTTGTGTGAGAGAGAGCATATTGAGAGGAGCATCTGTGTGTC 240  
QY 618 CTTGAAAGTATGATCAGAGAGGAGACCAAGGTGTGTTTCCCTTTGTGTTAAAGCTG 677  
DB 241 CTTGAAAGTATGATCAGAGAGGAGACCAAGGTGTGTTTCCCTTTGTGTTAAAGCTG 300  
QY 678 AGGCAAGAGGAGACCTTATCCAGCATTTTCCAAAGTGTGGGTGGTTCCTTCCGCA 737  
DB 301 AGGCAAGAGGAGACCTTATCCAGCATTTTCCAAAGTGTGGGTGGTTCCTTCCGCA 360  
QY 738 GATACCTTTAGTGTGATGGGCTGCTGATTAAGTGGCACAAAATCAGAGCAAGAGCA 797  
DB 361 GATACCTTTAGTGTGATGGGCTGCTGATTAAGTGGCACAAAATCAGAGCAAGAGCA 420  
QY 798 TGGCTTCCCAATTCCTCTCAATCTTTTATGCCGA 832  
DB 421 TGGCTTCCCAATTCCTCTCAATCTTTTATGCCGA 455

RESULT 13  
BG055477/c 446 bp mRNA linear EST 25-JAN-2001  
LOCUS nad5.shol.x1 NCI\_CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:3407472 3',  
DEFINITION mRNA sequence.  
ACCESSION BG055477  
VERSION BG055477.1 GI:12516626  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
REFERENCE 1 (bases 1 to 446)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLT, send email to:  
info@image.lnln.gov  
Seq primer: -400P from Gibco  
High quality sequence stop: 438.

FEATURES  
source  
Location/Qualifiers  
1..446  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:3407472"  
/issue\_type="carcinoid"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Lu24"  
/note="Organ: lung; Vector: pUT3D-Pac (Pharmacia) with a  
modified polylinker; Plasmid DNA from the normalized  
library NCI\_CGAP Lu5 was prepared, and ss circles were  
made in vitro. Following HAP purification, this DNA was  
used as tracer in a subtractive hybridization reaction.  
The driver was PCR-amplified cDNAs from a pool of 5,000  
clones made from the same library (clonids  
141920-1417991 and 1520904-1522439). Subtraction by Bento  
Soares and M. Fatima Bonaldo."

#### ORIGIN

Query Match 37.4%; Score 446; DB 10; Length 446;  
Best Local Similarity 100.0%; Pred. No. 1.2e-80;  
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 588 CAGGCAATTTTGGAGGGGATCTGTGTGCTTGAAGTATGATCAGAGAGGGAGCA 647  
DB 446 CAGGCAATTTTGGAGGGGATCTGTGTGCTTGAAGTATGATCAGAGAGGGAGCA 387  
QY 648 GGTGTGTGTTTCCCTTTGTTAGGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707  
DB 386 GGTGTGTGTTTCCCTTTGTTAGGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 327  
QY 708 CAAGGTGTGAGGTGCTGTTGTTCCGAGATACCTTTAGGTGTATGGGGCTGCATT 767  
DB 326 CAAGGTGTGAGGTGCTGTTGTTCCGAGATACCTTTAGGTGTATGGGGCTGCATT 267  
QY 768 AAGTGGCAAAATCAGAGCAAGAGAGAGATGCCCTTCCCAATCTCTCAATCCTTTAT 827  
DB 266 AAGTGGCAAAATCAGAGCAAGAGAGAGATGCCCTTCCCAATCTCTCAATCCTTTAT 207  
QY 828 GCCGAGAGATCTCAGCTGATGCAACATGTTCCGATGCTGTGAGAGAGATGCCGAG 887  
DB 206 GCCGAGAGATCTCAGCTGATGCAACATGTTCCGATGCTGTGAGAGAGATGCCGAG 147  
QY 888 TCTCTCTGCTTAGGAGACAGAGATTTGGGCTTAGGAGAGAGAGAGAGAGAGAGAGAG 947  
DB 146 TCTCTCTGCTTAGGAGACAGAGATTTGGGCTTAGGAGAGAGAGAGAGAGAGAGAGAG 87  
QY 948 TTAGCACTGTTTGTGTTTATAGTATATTTTATTTTATTTTATTTTATTTTATTTTATTT 1007  
DB 86 TTAGCACTGTTTGTGTTTATAGTATATTTTATTTTATTTTATTTTATTTTATTTTATTT 27  
QY 1008 GTGTAGTGCAATCTTATTTATTTG 1033  
DB 26 GTGTAGTGCAATCTTATTTATTTG 1

RESULT 14







ORIGIN

UT-E-C11 is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7m3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

TAG TISSUE=RPE and Choroid  
TAG\_LIB=UT-E-C11  
TAG\_SEQ=ACCTA"

Query Match 36.2%; Score 431; DB 13; Length 533;  
Best Local Similarity 99.6%; Pred. No. 1e-77;  
Matches 531; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

|    |      |                                                          |      |
|----|------|----------------------------------------------------------|------|
| QY | 544  | GGGTGGGTTTGTGAAGGACTTGTGGGCTTGGTGTGAAGCAGGATATTGAGG      | 603  |
| DB | 533  | GGGTGGGTTTGTGAAGGACTTGTGGGCTTGGTGTGAAGCAGGATATTGAGG      | 474  |
| QY | 604  | GGATCTGGTGGTCCCTTGAAGGATATAGAGAGGGGACACAGTGTGTCCCT       | 663  |
| DB | 473  | GGATCTGGTGGTCCCTTGAAGGATATAGAGAGGGGACACAGTGTGTCCCT       | 414  |
| QY | 664  | TTGTGTAAAGCGTGAAGGAGGAGACGTTAGTCCAGCATTTCCAAAGTGGGT      | 723  |
| DB | 413  | TTGTGTAAAGCGTGAAGGAGGAGACGTTAGTCCAGCATTTCCAAAGTGGGT      | 354  |
| QY | 724  | CGGTGGTCCCGAGTACTTTTAAAGTGTATGGGCTTGCACTTAAGTGGCAAAATCA  | 783  |
| DB | 353  | CGGTGGTCCCGAGTACTTTTAAAGTGTATGGGCTTGCACTTAAGTGGCAAAATCA  | 294  |
| QY | 784  | GAGCAGAAAGCGATGCCCTCCCAATTCCTCAATCCCTTTATGCCGAGAGATCTCAG | 843  |
| DB | 293  | GAGCAGAAAGCGATGCCCTCCCAATTCCTCAATCCCTTTATGCCGAGAGATCTCAG | 234  |
| QY | 844  | CTGGATGCCAAGTGTCCGATGCTGTGAAAGACATGCCGACGTCTCTGCTAGGG    | 903  |
| DB | 233  | CTGGATGCCAAGTGTCCGATGCTGTGAAAGACATGCCGACGTCTCTGCTAGGG    | 174  |
| QY | 904  | AGCAGACTTGGGCTTAGGGGAGGTGSAAAAATTCAGACTTTTGAAGCACTGTTTG  | 963  |
| DB | 173  | AGCAGACTTGGGCTTAGGGGAGGTGSAAAAATTCAGACTTTTGAAGCACTGTTTG  | 114  |
| QY | 964  | TTTAAATGATATTTTATTTGCTACTTTATTTGTTAGACAGTGTAGTGGCATTC    | 1023 |
| DB | 113  | TTTAAATGATATTTTATTTGCTACTTTATTTGTTAGACAGTGTAGTGGCATTC    | 54   |
| QY | 1024 | ATTATTTGACCTTTTCAATATAGATTAAATTAAGTAAAAAATTAAGTAAAAA     | 1076 |
| DB | 53   | ATTATTTGACCTTTTCAATATAGATTAAATTAAGTAAAAAATTAAGTAAAAA     | 1    |

Search completed: August 22, 2004, 07:09:17  
Job time : 3000 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: August 22, 2004, 04:42:00 ; Search time 117 Seconds

(without alignments)  
5653.859 Million cell updates/sec

Title: US-10-001-885-31

Sequence: 1 aaacgccttgagccgcgcg.....gaagagcacacggaacc 1192

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/2/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/prodata/2/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/prodata/2/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/prodata/2/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/prodata/2/ina/PCFUS.COMB.seq:\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match Length | ID                  | Description       |
|------------|-------|--------------------|---------------------|-------------------|
| 1          | 792   | 66.4               | US-09-149-476-58    | Sequence 58, Appl |
| 2          | 473   | 39.7               | US-09-621-976-864   | Sequence 864, App |
| 3          | 185   | 15.5               | US-09-621-976-3064  | Sequence 3064, Ap |
| 4          | 34    | 2.9                | US-08-747-221B-24   | Sequence 24, Appl |
| 5          | 34    | 2.9                | US-08-747-221B-26   | Sequence 26, Appl |
| 6          | 34    | 2.9                | US-09-005-051-24    | Sequence 24, Appl |
| 7          | 34    | 2.9                | US-09-005-051-26    | Sequence 26, Appl |
| 8          | 34    | 2.9                | US-09-403-942F-24   | Sequence 24, Appl |
| 9          | 34    | 2.9                | US-09-403-942F-26   | Sequence 26, Appl |
| 10         | 32    | 2.7                | US-09-621-976-17202 | Sequence 17202, A |
| 11         | 32    | 2.7                | US-09-621-976-17202 | Sequence 17202, A |
| 12         | 32    | 2.7                | US-08-368-776A-4    | Sequence 4, Appl  |
| 13         | 32    | 2.7                | PCT-US96-00419-4    | Sequence 4, Appl  |
| 14         | 32    | 2.7                | US-09-318-686-2     | Sequence 2, Appl  |
| 15         | 32    | 2.7                | US-09-318-686-1     | Sequence 1, Appl  |
| 16         | 31    | 2.6                | US-09-328-111-797   | Sequence 797, App |
| 17         | 31    | 2.6                | US-09-328-111-125   | Sequence 125, App |
| 18         | 31    | 2.6                | US-09-713-550-125   | Sequence 125, App |
| 19         | 31    | 2.6                | US-09-328-111-436   | Sequence 436, App |
| 20         | 31    | 2.6                | US-09-328-111-561   | Sequence 561, App |
| 21         | 31    | 2.6                | US-08-757-046A-5    | Sequence 5, Appl  |
| 22         | 31    | 2.6                | US-09-447-208-5     | Sequence 5, Appl  |
| 23         | 31    | 2.6                | US-09-135-988-5     | Sequence 5, Appl  |
| 24         | 31    | 2.6                | US-09-277-716-5     | Sequence 5, Appl  |
| 25         | 31    | 2.6                | US-08-577-274A-5    | Sequence 5, Appl  |
| 26         | 31    | 2.6                | US-08-908-909-5     | Sequence 5, Appl  |
| 27         | 31    | 2.6                | US-09-609-161B-5    | Sequence 5, Appl  |

|    |    |     |                    |                   |
|----|----|-----|--------------------|-------------------|
| 28 | 31 | 2.6 | US-08-990-103-5    | Sequence 5, Appl  |
| 29 | 31 | 2.6 | US-09-746-485A-5   | Sequence 5, Appl  |
| 30 | 31 | 2.6 | US-10-126-139-5    | Sequence 5, Appl  |
| 31 | 31 | 2.6 | US-10-126-798-5    | Sequence 5, Appl  |
| 32 | 31 | 2.6 | US-10-126-777-5    | Sequence 5, Appl  |
| 33 | 31 | 2.6 | US-09-426-557-1    | Sequence 1, Appl  |
| 34 | 31 | 2.6 | US-08-488-199-3    | Sequence 3, Appl  |
| 35 | 31 | 2.6 | US-08-147-777-3    | Sequence 3, Appl  |
| 36 | 31 | 2.6 | US-08-452-872-3    | Sequence 3, Appl  |
| 37 | 31 | 2.6 | PCT-US93-03985-3   | Sequence 3, Appl  |
| 38 | 31 | 2.6 | US-09-497-855A-37  | Sequence 37, Appl |
| 39 | 31 | 2.6 | US-09-497-855A-44  | Sequence 44, Appl |
| 40 | 30 | 2.5 | US-09-197-814-9    | Sequence 9, Appl  |
| 41 | 30 | 2.5 | US-09-320-581-9    | Sequence 9, Appl  |
| 42 | 30 | 2.5 | US-09-165-239A-5   | Sequence 5, Appl  |
| 43 | 30 | 2.5 | US-09-894-916-3    | Sequence 3, Appl  |
| 44 | 30 | 2.5 | US-09-404-879A-201 | Sequence 201, App |
| 45 | 30 | 2.5 | US-09-338-933-201  | Sequence 201, App |

#### ALIGNMENTS

RESULT 1  
US-09-149-476-58  
Sequence 58, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION: a1.  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
EARLIER FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23

|    |                                        |    |                                        |
|----|----------------------------------------|----|----------------------------------------|
| 1  | EARLIER APPLICATION NUMBER: 60/047,500 | 1  | EARLIER FILING DATE: 1997-08-22        |
| 2  | EARLIER FILING DATE: 1997-05-23        | 2  | EARLIER APPLICATION NUMBER: 60/056,899 |
| 3  | EARLIER APPLICATION NUMBER: 60/047,587 | 3  | EARLIER FILING DATE: 1997-08-22        |
| 4  | EARLIER FILING DATE: 1997-05-23        | 4  | EARLIER APPLICATION NUMBER: 60/056,911 |
| 5  | EARLIER APPLICATION NUMBER: 60/047,492 | 5  | EARLIER FILING DATE: 1997-08-22        |
| 6  | EARLIER FILING DATE: 1997-05-23        | 6  | EARLIER APPLICATION NUMBER: 60/056,633 |
| 7  | EARLIER APPLICATION NUMBER: 60/047,598 | 7  | EARLIER FILING DATE: 1997-08-22        |
| 8  | EARLIER FILING DATE: 1997-05-23        | 8  | EARLIER APPLICATION NUMBER: 60/056,877 |
| 9  | EARLIER APPLICATION NUMBER: 60/047,613 | 9  | EARLIER FILING DATE: 1997-08-22        |
| 10 | EARLIER FILING DATE: 1997-05-23        | 10 | EARLIER APPLICATION NUMBER: 60/056,910 |
| 11 | EARLIER APPLICATION NUMBER: 60/047,582 | 11 | EARLIER FILING DATE: 1997-08-22        |
| 12 | EARLIER FILING DATE: 1997-05-23        | 12 | EARLIER APPLICATION NUMBER: 60/056,866 |
| 13 | EARLIER APPLICATION NUMBER: 60/047,596 | 13 | EARLIER FILING DATE: 1997-08-22        |
| 14 | EARLIER FILING DATE: 1997-05-23        | 14 | EARLIER APPLICATION NUMBER: 60/056,633 |
| 15 | EARLIER APPLICATION NUMBER: 60/047,612 | 15 | EARLIER FILING DATE: 1997-08-22        |
| 16 | EARLIER FILING DATE: 1997-05-23        | 16 | EARLIER APPLICATION NUMBER: 60/056,845 |
| 17 | EARLIER APPLICATION NUMBER: 60/047,632 | 17 | EARLIER FILING DATE: 1997-08-22        |
| 18 | EARLIER FILING DATE: 1997-05-23        | 18 | EARLIER APPLICATION NUMBER: 60/056,899 |
| 19 | EARLIER APPLICATION NUMBER: 60/047,601 | 19 | EARLIER FILING DATE: 1997-08-22        |
| 20 | EARLIER FILING DATE: 1997-05-23        | 20 | EARLIER APPLICATION NUMBER: 60/057,764 |
| 21 | EARLIER APPLICATION NUMBER: 60/043,580 | 21 | EARLIER FILING DATE: 1997-08-22        |
| 22 | EARLIER FILING DATE: 1997-04-11        | 22 | EARLIER APPLICATION NUMBER: 60/047,599 |
| 23 | EARLIER APPLICATION NUMBER: 60/043,568 | 23 | EARLIER FILING DATE: 1997-05-23        |
| 24 | EARLIER FILING DATE: 1997-04-11        | 24 | EARLIER APPLICATION NUMBER: 60/047,599 |
| 25 | EARLIER APPLICATION NUMBER: 60/043,314 | 25 | EARLIER FILING DATE: 1997-05-23        |
| 26 | EARLIER FILING DATE: 1997-04-11        | 26 | EARLIER APPLICATION NUMBER: 60/047,588 |
| 27 | EARLIER APPLICATION NUMBER: 60/043,569 | 27 | EARLIER FILING DATE: 1997-05-23        |
| 28 | EARLIER FILING DATE: 1997-04-11        | 28 | EARLIER APPLICATION NUMBER: 60/047,588 |
| 29 | EARLIER APPLICATION NUMBER: 60/043,311 | 29 | EARLIER FILING DATE: 1997-05-23        |
| 30 | EARLIER FILING DATE: 1997-04-11        | 30 | EARLIER APPLICATION NUMBER: 60/047,588 |
| 31 | EARLIER APPLICATION NUMBER: 60/043,671 | 31 | EARLIER FILING DATE: 1997-05-23        |
| 32 | EARLIER FILING DATE: 1997-04-11        | 32 | EARLIER APPLICATION NUMBER: 60/047,586 |
| 33 | EARLIER APPLICATION NUMBER: 60/043,674 | 33 | EARLIER FILING DATE: 1997-05-23        |
| 34 | EARLIER FILING DATE: 1997-04-11        | 34 | EARLIER APPLICATION NUMBER: 60/047,590 |
| 35 | EARLIER APPLICATION NUMBER: 60/043,669 | 35 | EARLIER FILING DATE: 1997-05-23        |
| 36 | EARLIER FILING DATE: 1997-04-11        | 36 | EARLIER APPLICATION NUMBER: 60/047,599 |
| 37 | EARLIER APPLICATION NUMBER: 60/043,312 | 37 | EARLIER FILING DATE: 1997-05-23        |
| 38 | EARLIER FILING DATE: 1997-04-11        | 38 | EARLIER APPLICATION NUMBER: 60/047,593 |
| 39 | EARLIER APPLICATION NUMBER: 60/043,313 | 39 | EARLIER FILING DATE: 1997-05-23        |
| 40 | EARLIER FILING DATE: 1997-04-11        | 40 | EARLIER APPLICATION NUMBER: 60/047,614 |
| 41 | EARLIER APPLICATION NUMBER: 60/043,672 | 41 | EARLIER FILING DATE: 1997-05-23        |
| 42 | EARLIER FILING DATE: 1997-04-11        | 42 | EARLIER APPLICATION NUMBER: 60/043,578 |
| 43 | EARLIER APPLICATION NUMBER: 60/043,315 | 43 | EARLIER FILING DATE: 1997-04-11        |
| 44 | EARLIER FILING DATE: 1997-04-11        | 44 | EARLIER APPLICATION NUMBER: 60/043,576 |
| 45 | EARLIER APPLICATION NUMBER: 60/048,974 | 45 | EARLIER FILING DATE: 1997-04-11        |
| 46 | EARLIER FILING DATE: 1997-06-06        | 46 | EARLIER APPLICATION NUMBER: 60/047,501 |
| 47 | EARLIER APPLICATION NUMBER: 60/056,886 | 47 | EARLIER FILING DATE: 1997-05-23        |
| 48 | EARLIER FILING DATE: 1997-08-22        | 48 | EARLIER APPLICATION NUMBER: 60/043,670 |
| 49 | EARLIER APPLICATION NUMBER: 60/056,877 | 49 | EARLIER FILING DATE: 1997-04-11        |
| 50 | EARLIER FILING DATE: 1997-08-22        | 50 | EARLIER APPLICATION NUMBER: 60/056,632 |
| 51 | EARLIER APPLICATION NUMBER: 60/056,889 | 51 | EARLIER FILING DATE: 1997-08-22        |
| 52 | EARLIER FILING DATE: 1997-08-22        | 52 | EARLIER APPLICATION NUMBER: 60/056,664 |
| 53 | EARLIER APPLICATION NUMBER: 60/056,893 | 53 | EARLIER FILING DATE: 1997-08-22        |
| 54 | EARLIER FILING DATE: 1997-08-22        | 54 | EARLIER APPLICATION NUMBER: 60/056,876 |
| 55 | EARLIER APPLICATION NUMBER: 60/056,630 | 55 | EARLIER FILING DATE: 1997-08-22        |
| 56 | EARLIER FILING DATE: 1997-08-22        | 56 | EARLIER APPLICATION NUMBER: 60/056,881 |
| 57 | EARLIER APPLICATION NUMBER: 60/056,878 | 57 | EARLIER FILING DATE: 1997-08-22        |
| 58 | EARLIER FILING DATE: 1997-08-22        | 58 | EARLIER APPLICATION NUMBER: 60/056,909 |
| 59 | EARLIER APPLICATION NUMBER: 60/056,662 | 59 | EARLIER FILING DATE: 1997-08-22        |
| 60 | EARLIER FILING DATE: 1997-08-22        | 60 | EARLIER APPLICATION NUMBER: 60/056,875 |
| 61 | EARLIER APPLICATION NUMBER: 60/056,872 | 61 | EARLIER FILING DATE: 1997-08-22        |
| 62 | EARLIER FILING DATE: 1997-08-22        | 62 | EARLIER APPLICATION NUMBER: 60/056,862 |
| 63 | EARLIER APPLICATION NUMBER: 60/056,882 | 63 | EARLIER FILING DATE: 1997-08-22        |
| 64 | EARLIER FILING DATE: 1997-08-22        | 64 | EARLIER APPLICATION NUMBER: 60/056,887 |
| 65 | EARLIER APPLICATION NUMBER: 60/056,637 | 65 | EARLIER FILING DATE: 1997-08-22        |
| 66 | EARLIER FILING DATE                    |    |                                        |

EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 56.4%; Score 792; DB 4; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0; Mismatches 0; Indels 2; Gaps 2;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

46 TCGCCCTGACAGACAGATCTACTAGCGGTGTCACCTTGTGAACATCATCTAGCTG 105  
1 TCGCCCTGACAGACAGATCTACTAGCGGTGTCACCTTGTGAACATCATCTAGCTG 60  
106 CAAGCTCTCTCAATTTCTGTGAGCCCATCAGGAGCCCAAGGCGCTGGAGAGATG 165  
61 CAAGCTCTCTCAATTTCTGTGAGCCCATCAGGAGCCCAAGGCGCTGGAGAGATG 120  
166 GCGGATCTTGCCCAATGAGGGAATGCTGACAGACAGACCCCGAGTGAAGACCATAC 225  
121 GCGGATCTTGCCCAATGAGGGAATGCTGACAGACAGACCCCGAGTGAAGACCATAC 180  
226 CCAAGCCACCAAGATGATCTCTGACAGAGCTTTTCAATAGGGGCAATGATGCTCC 285  
181 CCAAGCCACCAAGATGATCTCTGACAGAGCTTTTCAATAGGGGCAATGATGCTCC 240  
286 CCAAGGCGGTCTGCGCCCGCCGACAGACAGGACAGGTCAGAGTGGTGTCTCAGTC 345  
241 CCAAGGCGGTCTGCGCCCGCCGACAGACAGGACAGGTCAGAGTGGTGTCTCAGTC 300  
346 CCGCTTCAATGACCTCAACCGGACGCTGTGAACATGAGCTTTCCGACATGCTCGG 405  
301 CCGCTTCAATGACCTCAACCGGACGCTGTGAACATGAGCTTTCCGACATGCTCGG 360  
406 CAACCATGCTGTGAGACCGGTGACCTGCACTCGTCCCTTCTGCTCATGATGCTGG 465  
361 CAACCATGCTGTGAGACCGGTGACCTGCACTCGTCCCTTCTGCTCATGATGCTGG 420  
466 TGTTCGTGCGCTCTCTGCTGTGAGCTTTGCTAAGCTGTGCTCCAGCTGAGCTG 525  
421 TGTTCGTGCGCTCTCTGCTGTGAGCTTTGCTAAGCTGTGCTCCAGCTGAGCTG 480  
526 ACCTCTGAGGCGTGAATGAGGCGGTGTTGTTGAGAGGACTTGTGAGGCGCTG 585  
481 ACCTCTGAGGCGTGAATGAGGCGGTGTTGTTGAGAGGACTTGTGAGGCGCTG 540  
586 AGCAGGCAATTTGAGAGGAGATGCTGAGGCTTGAAGGATGATCAGAGAGGAGCA 645  
541 AGCAGGCAATTTGAGAGGAGATGCTGAGGCTTGAAGGATGATCAGAGAGGAGCA 600  
646 CAGGTGTGTTTCCCTTTGTGTTAAGCTGAGGCAAGAGAGCTTAACT-CCAGCAT 704  
601 CAGGTGTGTTTCCCTTTGTGTTAAGCTGAGGCAAGAGAGCTTAACT-CCAGCAT 660  
705 TT-CCAAAGTGTGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763  
661 TTCCCAAGTGTGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
764 CATTAAAGTGCACAAATCAGAGCAAGAAAGCATGCTTCCCAATTTCTCAATCTT 823  
721 CATTAAAGTGCACAAATCAGAGCAAGAAAGCATGCTTCCCAATTTCTCAATCTT 780  
824 TTATGCGAGAAATCTCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883  
781 TTATGCGAGAAATCTCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
884 GAGGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943  
841 GAGGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
944 CTTTGTAGCACTGTTTGTGTTTAAATGATATTTTATGCTTATGTTAGG 1003

901 CTTTGTAGCACTGTTTGTGTTTAAATGATATTTTATGCTTATGTTAGG 960  
1004 ACAAGTGTAGTGGCATTTCTATTTATTTGACCTTTTCAATAATGATTTAGTAAAA 1063  
961 ACAAGTGTAGTGGCATTTCTATTTATTTGACCTTTTCAATAATGATTTAGTAAAA 1020  
1064 AAAAAAAAAAAAAA 1077  
1021 AAAAAAAAAAAAAA 1034

RESULT 2  
US-09-621-976-864  
Sequence 864, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OR INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent pm  
SEQ ID NO 864  
LENGTH: 483  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURES:  
NAME/KEY: CDS  
LOCATION: 171..482  
US-09-621-976-864

Query Match 39.7%; Score 473; DB 4; Length 483;  
Best Local Similarity 100.0%; Pred. No. 9,6e-202; Indels 0; Gaps 0;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAACGGCTGAGGCGCGGAGGTGAGACGCGCCGAGCCGAGTGGCGCTGACAGCA 60  
11 AAACGGCTGAGGCGCGGAGGTGAGACGCGCCGAGCCGAGTGGCGCTGACAGCA 70  
61 CAGCATCTACTCAGCGTGGGTCACTCTGTGACATCACTGACTGCAAGCTTCCCTCAAT 120  
71 CAGCATCTACTCAGCGTGGGTCACTCTGTGACATCACTGACTGCAAGCTTCCCTCAAT 130  
121 TTCTGTGACAGCCATCAGGAGCCCAAGCGCTGAGAGATGATGAGGATCTTGCCCA 180  
131 TTCTGTGACAGCCATCAGGAGCCCAAGCGCTGAGAGATGATGAGGATCTTGCCCA 190  
181 TGGGGAATCTGTGAGAGAGAGACCCCGAGTGAAGCACTAACCCAGCAAGAGG 240  
191 TGGGGAATCTGTGAGAGAGAGACCCCGAGTGAAGCACTAACCCAGCAAGAGG 250  
241 TAGCATCTCTGACAGAGCTTCTTAATAGAGGCAATGATGCTTCCCGAGGCGCTG 300  
251 TAGCATCTCTGACAGAGCTTCTTAATAGAGGCAATGATGCTTCCCGAGGCGCTG 310  
301 CCGCGCGAGAGAGAGAGGAGGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTG 360  
311 CCGCGCGAGAGAGAGAGGAGGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTG 370  
361 CAACCGGAGAGGAGGAGGAGGCTTCCGAGTGAATCTTCCGCAACCATGCTGTGA 420  
371 CAACCGGAGAGGAGGAGGAGGCTTCCGAGTGAATCTTCCGCAACCATGCTGTGA 430  
421 GCGGTGACCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 473  
431 GCGGTGACCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483

RESULT 3  
US-09-621-976-3064

Tue Aug 24 09:48:32 2004

us-10-001-885-31.011.rn1

Page 4

Sequence 3064, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
FILE REFERENCE: GENSET 054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.Pm  
SEQ ID NO 3064  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 93..431  
US-09-621-976-3064

Query Match 15.5%; Score 185; DB 4; Length 519;  
Best Local Similarity 99.6%; Pred. No. 3.2e-73;  
Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 ACAAGCATCTACTCAGCGTGGTCACTCTGTGAACATCACTGACGAGGCTCCCTC 117  
DB 116 ACACAGCATCTACTCAGCGTGGTCACTCTGTGAACATCACTGACGAGGCTCCCTC 175  
QY 118 AATTTCGTGTCAGCCCATAGGACCCACAGCGCTCGAGAGATGTCGATCTTGGC 177  
DB 176 AATTTCGTGTCAGCCCATAGGACCCACAGCGCTCGAGAGATGTCGATCTTGGC 235  
QY 178 CAATGGGGAATGTCGAGGACGACGACCCCGAGAGAGACCACTACACCCACCCACG 237  
DB 236 CAATGGGGAATGTCGAGGACGACGACCCCGAGAGAGACCACTACACCCACCCACG 295  
QY 238 AGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGGCAATGTCGTCCTCCCAAGG 293  
DB 296 AGGTAGCATCTCTGACAGAGCTTCTTCAATAGGGGCAATGTCGTCCTCCCAAGG 351

RESULT 4

US-08-747-221B-24  
Sequence 24, Application US/08747221B  
Patent No. 6063610

GENERAL INFORMATION:  
APPLICANT: Silver, Gary W.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: No. 6063610e1 Carboxylesterase Nucleic Acid  
TITLE OF INVENTION: Molecules, Proteins and Uses Thereof  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/747,221B  
FILING DATE: No. 6063610e1 October 12, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: FC-1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2836 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 99..1889  
US-08-747-221B-24

Query Match 2.9%; Score 34; DB 3; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 7.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TTTAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1085  
DB 2778 TTTAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2811

RESULT 5

US-08-747-221B-26/C  
Sequence 26, Application US/08747221B  
Patent No. 6063610

GENERAL INFORMATION:  
APPLICANT: Silver, Gary W.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: No. 6063610e1 Carboxylesterase Nucleic Acid  
TITLE OF INVENTION: Molecules, Proteins and Uses Thereof  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/747,221B  
FILING DATE: No. 6063610e1 October 12, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: FC-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2836 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-747-221B-26

Query Match 2.9%; Score 34; DB 3; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 7.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TTTAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1085

Db 59 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 26

RESULT 6  
US-09-005-051-24  
Sequence 24, Application US/09005051

PATENT INFORMATION:

APPLICANT: Silver, Gary W.

APPLICANT: Wisniewski, Nancy

TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid

NUMBER OF SEQUENCES: 66

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carol Talkington Verser, Ph.D.

ADDRESSEE: Heska Corporation

STREET: 1825 Sharp Point Drive

CITY: Fort Collins

STATE: Colorado

COUNTRY: USA

ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: Wordperfect for Windows, Version 7.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/005,051

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/747,221

FILING DATE: NO. 6291222el December 12, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington

REGISTRATION NUMBER: 37,459

REFERENCE/DOCKET NUMBER: FC-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 970/493-7272

TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 2836 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 99..1889

US-09-005-051-24

Query Match 2.9%; Score 34; DB 3; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 7.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 1085  
2778 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 2811

RESULT 7  
US-09-005-051-26/c  
Sequence 26, Application US/09005051

PATENT INFORMATION:

APPLICANT: Silver, Gary W.

APPLICANT: Wisniewski, Nancy

TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid

NUMBER OF SEQUENCES: 66

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carol Talkington Verser, Ph.D.

ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: Wordperfect for Windows, Version 7.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/005,051

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/747,221

FILING DATE: NO. 6291222el December 12, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington

REGISTRATION NUMBER: 37,459

REFERENCE/DOCKET NUMBER: FC-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 970/493-7272

TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 2836 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-005-051-26

Query Match 2.9%; Score 34; DB 3; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 7.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 1085  
59 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 26

RESULT 8  
US-09-403-942F-24  
Sequence 24, Application US/09403942F

PATENT INFORMATION:

APPLICANT: Silver, Gary W.

APPLICANT: Brandt, Kevin S.

TITLE OF INVENTION: NOVEL CARBOXYLESTERASE NUCLEIC ACID MOLECULES, PROTEINS AND USES

FILE REFERENCE: FC-1-CI-PUS

CURRENT APPLICATION NUMBER: US/09/403,942F

PRIOR FILING DATE: 2000-05-02

PRIOR APPLICATION NUMBER: PCT/US97/20598

PRIOR FILING DATE: 1997-11-10

PRIOR APPLICATION NUMBER: 08/747,221

PRIOR FILING DATE: 1996-11-12

NUMBER OF SEQ ID NOS: 76

SOFTWARE: Patent version 3.1

SEQ ID NO 24

LENGTH: 2836

TYPE: DNA

ORGANISM: Ctenocephalides felis

FEATURE:

NAME/KEY: CDS

LOCATION: (99)..(1889)

OTHER INFORMATION:

FEATURE: misc\_feature

NAME/KEY: misc\_feature

LOCATION: (2278)..(2278)

OTHER INFORMATION: n = unknown





REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2901 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic  
US-08-368-776A-4

Query Match 2.7%; Score 32; DB 4; Length 2901;  
Best Local Similarity 100.0%; Pred. No. 5.8e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 TAAGTAAAAAAAAAAAAAAAAAAAAAAAAA 1085  
DB 2853 TAAGTAAAAAAAAAAAAAAAAAAAAAAAAA 2884

RESULT 13  
PCT-US96-00419-4  
Sequence 4, Application PC/TUS9600419  
GENERAL INFORMATION:  
APPLICANT: Thomas Closek, Axel Ullrich, Birgit  
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF  
TITLE OF INVENTION: MDX1 SIGNAL TRANSDUCTION DISORDERS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/00419  
FILING DATE: January 3, 1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: none  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2901  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic  
PCT-US96-00419-4

Query Match 2.7%; Score 32; DB 5; Length 2901;  
Best Local Similarity 100.0%; Pred. No. 5.8e-05;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1054 TAAGTAAAAAAAAAAAAAAAAAAAAAAAAA 1085  
DB 2853 TAAGTAAAAAAAAAAAAAAAAAAAAAAAAA 2884

RESULT 14  
US-09-918-686-2  
Sequence 2, Application US/09918686  
Patent No. 6475739  
GENERAL INFORMATION:  
APPLICANT: Brunkow, Mary  
APPLICANT: Prohl, Sean  
APPLICANT: Paepfer, Bryan  
APPLICANT: Staehling-Hampton, Karen  
TITLE OF INVENTION: METHODS FOR IDENTIFYING  
TITLE OF INVENTION: GENOMIC DELETIONS  
FILE REFERENCE: 240083.515  
CURRENT APPLICATION NUMBER: US/09/918, 686  
CURRENT FILING DATE: 2001-07-30  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 51719  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1246..2572, 2604  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-686-2

Query Match 2.7%; Score 32; DB 4; Length 51719;  
Best Local Similarity 100.0%; Pred. No. 5.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAG 1086  
DB 44873 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAG 44904

RESULT 15  
US-09-918-686-1  
Sequence 1, Application US/09918686  
Patent No. 6475739  
GENERAL INFORMATION:  
APPLICANT: Brunkow, Mary  
APPLICANT: Prohl, Sean  
APPLICANT: Paepfer, Bryan  
APPLICANT: Staehling-Hampton, Karen  
TITLE OF INVENTION: METHODS FOR IDENTIFYING  
TITLE OF INVENTION: GENOMIC DELETIONS  
FILE REFERENCE: 240083.515  
CURRENT APPLICATION NUMBER: US/09/918, 686  
CURRENT FILING DATE: 2001-07-30  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 92139  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 7043..8369, 8401  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-686-1

Query Match 2.7%; Score 32; DB 4; Length 92139;  
Best Local Similarity 100.0%; Pred. No. 5.2e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAG 1086

Tue Aug 24 09:48:32 2004

us-10-001-885-31.014.rml

Page 8

Db 50660 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 50691

Search completed: August 22, 2004, 07:11:24  
Job time : 119 secs

OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 06:19:16 ; Search time 594 Seconds  
(without alignments)  
9857.345 Million cell updates/sec

Title: US-10-001-885-31

Perfect score: 1192

Sequence: 1 aaacgcctgagagccgcgcg.....gaagagagacacggaacc 1192

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3228839 seqs, 245606551 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 1192  | 100.0       | 1192   | 13    | US-10-001-885-31     |
| 2          | 903   | 75.8        | 1108   | 16    | US-10-264-237-619    |
| 3          | 792   | 66.4        | 1049   | 10    | US-09-809-331-58     |
| 4          | 792   | 66.4        | 1049   | 10    | US-09-882-171-58     |
| 5          | 792   | 66.4        | 1049   | 13    | US-10-164-861-94     |
| 6          | 669   | 56.1        | 1052   | 13    | US-10-302-172-94     |
| 7          | 326   | 27.3        | 1052   | 15    | US-10-076-747-65     |
| 8          | 195   | 16.4        | 612    | 15    | US-10-076-747-64     |
| 9          | 125   | 10.5        | 643    | 13    | US-10-001-885-30     |
| 10         | 38    | 3.2         | 473    | 9     | US-09-864-761-551    |
| 11         | 38    | 3.2         | 473    | 9     | US-09-864-761-551    |
| 12         | 38    | 3.2         | 998    | 13    | US-10-027-632-209023 |
| 13         | 38    | 3.2         | 998    | 16    | US-10-027-632-209023 |
| 14         | 34    | 2.9         | 616    | 13    | US-10-424-599-125925 |

|    |    |     |         |    |                      |                       |
|----|----|-----|---------|----|----------------------|-----------------------|
| 15 | 34 | 2.9 | 2836    | 17 | US-10-678-521-24     | Sequence 24, Appl     |
| 16 | 34 | 2.9 | 2836    | 17 | US-10-678-521-26     | Sequence 26, Appl     |
| 17 | 33 | 2.8 | 653     | 12 | US-09-969-034-718    | Sequence 718, Appl    |
| 18 | 33 | 2.8 | 758     | 13 | US-10-027-632-127811 | Sequence 127811, Appl |
| 19 | 33 | 2.8 | 758     | 16 | US-10-027-632-127811 | Sequence 127811, Appl |
| 20 | 33 | 2.8 | 833     | 15 | US-10-198-846-5140   | Sequence 5140, Appl   |
| 21 | 33 | 2.8 | 1473    | 17 | US-10-437-863-99876  | Sequence 99876, Appl  |
| 22 | 33 | 2.8 | 1885    | 16 | US-10-574-780A-849   | Sequence 849, Appl    |
| 23 | 33 | 2.8 | 96595   | 12 | US-09-997-722-202    | Sequence 202, Appl    |
| 24 | 33 | 2.8 | 21231   | 13 | US-10-087-192-1116   | Sequence 1126, Appl   |
| 25 | 33 | 2.8 | 3673778 | 15 | US-10-312-841-2      | Sequence 2, Appl      |
| 26 | 32 | 2.7 | 152     | 9  | US-09-834-975-110    | Sequence 110, Appl    |
| 27 | 32 | 2.7 | 157     | 9  | US-09-867-701-9937   | Sequence 9937, Appl   |
| 28 | 32 | 2.7 | 389     | 17 | US-10-437-863-26389  | Sequence 26389, Appl  |
| 29 | 32 | 2.7 | 473     | 16 | US-10-060-036-916    | Sequence 916, Appl    |
| 30 | 32 | 2.7 | 480     | 16 | US-10-131-827-8192   | Sequence 8192, Appl   |
| 31 | 32 | 2.7 | 572     | 17 | US-10-437-963-88429  | Sequence 88429, Appl  |
| 32 | 32 | 2.7 | 602     | 9  | US-09-918-686-7      | Sequence 7, Appl      |
| 33 | 32 | 2.7 | 602     | 15 | US-10-353-150-7      | Sequence 7, Appl      |
| 34 | 32 | 2.7 | 649     | 15 | US-10-066-443-727    | Sequence 727, Appl    |
| 35 | 32 | 2.7 | 706     | 15 | US-10-050-704-62     | Sequence 62, Appl     |
| 36 | 32 | 2.7 | 706     | 17 | US-10-798-512-62     | Sequence 62, Appl     |
| 37 | 32 | 2.7 | 1525    | 9  | US-09-860-107-3816   | Sequence 3816, Appl   |
| 38 | 32 | 2.7 | 1852    | 17 | US-10-437-963-5371   | Sequence 5371, Appl   |
| 39 | 32 | 2.7 | 2138    | 10 | US-09-814-353-19280  | Sequence 19280, Appl  |
| 40 | 32 | 2.7 | 3031    | 13 | US-10-027-632-113397 | Sequence 113397, Appl |
| 41 | 32 | 2.7 | 3031    | 16 | US-10-027-632-113397 | Sequence 113397, Appl |
| 42 | 32 | 2.7 | 3685    | 9  | US-09-822-830A-331   | Sequence 331, Appl    |
| 43 | 32 | 2.7 | 5204    | 15 | US-10-311-455-873    | Sequence 873, Appl    |
| 44 | 32 | 2.7 | 33112   | 16 | US-10-429-873A-3     | Sequence 3, Appl      |
| 45 | 32 | 2.7 | 39001   | 16 | US-10-199-199-14     | Sequence 14, Appl     |

#### ALIGNMENTS

RESULT 1  
US-10-001-885-31  
Sequence 31, Application US/10001885  
Publication No US20040058319A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Recipon, Hervé  
APPLICANT: Caffarely, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro  
FILE REFERENCE: DEX-0279  
CURRENT APPLICATION NUMBER: US/10/001,885  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/252,061  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: 60/253,257  
NUMBER OF SEQ ID NOS: 167  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 31  
LENGTH: 1192  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-001-885-31

Query Match 100.0%; Score 1192; DB 13; Length 1192;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAACGCTGAGAGCCGCGGAGTGAACCGCCGAGCCCGGAGTGCCTGCTGACACA 60  
1 AAACGCTGAGAGCCGCGGAGTGAACCGCCGAGCCCGGAGTGCCTGCTGACACA 60  
CAGCATCTACCTGAGTGTGTCACCTCTGTAACATGACTGACAGGCTCCCTCAAT 120

Db 61 CAGATCTACTCAGCGGTGCTCACTCTGTGAACATCACTGACTGCAAGCTCCCTCAAT 120  
QY 121 TTCTGTGCGAGCCCATCAGAGAACCCACAGGCTCTGGAGAGATGTCGGACCA 180  
Db 121 TTCTGTGCGAGCCCATCAGAGAACCCACAGGCTCTGGAGAGATGTCGGACCA 180  
QY 181 TGGGGAATATGTGACAGACAGACACCCCGATGAGACCACTACCCAGCACAAGAG 240  
Db 181 TGGGGAATATGTGACAGACAGACACCCCGATGAGACCACTACCCAGCACAAGAG 240  
QY 241 TAGCATCTCTGACAGAGCTTTCTCAATAGGGCCATGTCCTCCCGAGGGGCTCTGG 300  
Db 241 TAGCATCTCTGACAGAGCTTTCTCAATAGGGCCATGTCCTCCCGAGGGGCTCTGG 300  
QY 301 CCCCCCGACAGACAGAGAGGTGCAAGGCTGGGCTGCTGAGTCCCTTCAATGACT 360  
Db 301 CCCCCCGACAGACAGAGAGGTGCAAGGCTGGGCTGCTGAGTCCCTTCAATGACT 360  
QY 361 CAACCCGACAGCTGTGAACATGAGGCTTTCCGAGTGGCAATCTCGCAACCATGCTGTGA 420  
Db 361 CAACCCGACAGCTGTGAACATGAGGCTTTCCGAGTGGCAATCTCGCAACCATGCTGTGA 420  
QY 421 GCGGTGACCTCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
Db 421 GCGGTGACCTCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
QY 481 CCTGTGTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
Db 481 CCTGTGTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
QY 541 TAGGGGTGGGTTGTGTGAAGGACCTTGTGAGGACCTTGTGAGAGCAGCATATTGG 600  
Db 541 TAGGGGTGGGTTGTGTGAAGGACCTTGTGAGGACCTTGTGAGAGCAGCATATTGG 600  
QY 601 AGGGGATCTGTGTGTGAAGATGATCAGAGAGGACCAAGAGTGTGTGTTC 660  
Db 601 AGGGGATCTGTGTGTGAAGATGATCAGAGAGGACCAAGAGTGTGTGTTC 660  
QY 661 CCTTTGTGAAGCTGAGGACAGAGGACGTTAGTCCAGCATTTCCAAAGTGTGGTG 720  
Db 661 CCTTTGTGAAGCTGAGGACAGAGGACGTTAGTCCAGCATTTCCAAAGTGTGGTG 720  
QY 721 GGTGCTGTGCTCCGAGATCACTTATAGGTATGAGGAGCTGCAATTAAGTGCACAAA 780  
Db 721 GGTGCTGTGCTCCGAGATCACTTATAGGTATGAGGAGCTGCAATTAAGTGCACAAA 780  
QY 781 TCAAGACAGAAAGCGATGCTCCCAATCTCTCAATCTTTTATGCCAGAAAGATCT 840  
Db 781 TCAAGACAGAAAGCGATGCTCCCAATCTCTCAATCTTTTATGCCAGAAAGATCT 840  
QY 841 CAGCTGATGACCAAGATGTCGATGCTGTGAGACATGCCAGATCTCTCTGCTCA 900  
Db 841 CAGCTGATGACCAAGATGTCGATGCTGTGAGACATGCCAGATCTCTCTGCTCA 900  
QY 901 GAGAGCAGAGCTGTGAGGCTAGGCAAGTGAAGAAATTCAGACTTTTATAGCATGTT 960  
Db 901 GAGAGCAGAGCTGTGAGGCTAGGCAAGTGAAGAAATTCAGACTTTTATAGCATGTT 960  
QY 961 TTGTTTAAATGATATTTTATATGCTATTTATTTTATGAGCAAGTGTGTGCAAT 1020  
Db 961 TTGTTTAAATGATATTTTATATGCTATTTATTTTATGAGCAAGTGTGTGCAAT 1020  
QY 1021 TCAATTAATGATGATTTTCAATTAATGATTTAATGATTTAATGATTTAATGATTTA 1080  
Db 1021 TCAATTAATGATGATTTTCAATTAATGATTTAATGATTTAATGATTTAATGATTTA 1080  
QY 1081 AAAAAGCTGTGAGGCTTACCCGAGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
Db 1081 AAAAAGCTGTGAGGCTTACCCGAGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
QY 1141 GCGCCCAATTTCCGCAATTTTGTGACAAATGAAAGAGACACGGAACCC 1192  
Db 1141 GCGCCCAATTTCCGCAATTTTGTGACAAATGAAAGAGACACGGAACCC 1192  
QY 1141 GCGCCCAATTTCCGCAATTTTGTGACAAATGAAAGAGACACGGAACCC 1192  
Db 1141 GCGCCCAATTTCCGCAATTTTGTGACAAATGAAAGAGACACGGAACCC 1192

RESULT 2  
US-10-264-237-619  
; Sequence 619, Application US/10264237  
; Publication No. US2004009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birex et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P131p1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; PRIOR FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 619  
; LENGTH: 1108  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURES:  
; NAME/KEY: misc feature  
; LOCATION: (1101)..(1101)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-264-237-619  
Query Match 75.8%; Score 903; DB 16; Length 1108;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1053; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 22 AGTGAAGCGCCCGAGGCGCGAGATGCGGCTGTGAGACAGACATCACTCAGCTGGGT 81  
Db 8 AGTGAAGCGCCCGAGGCGCGAGATGCGGCTGTGAGACAGACATCACTCAGCTGGGT 67  
QY 82 CACCTGTGAACATCACTGACCTGACAGCTCCCTCAATTTCTGTGTCAGACCCATCAGG 141  
Db 68 CACCTGTGAACATCACTGACCTGACAGCTCCCTCAATTTCTGTGTCAGACCCATCAGG 127  
QY 142 ACCCAAGCGCTGGAGAGATGTCGAGATCTTGGCCCAATGGGAAATGTCGAGAGCA 201  
Db 128 ACCCAAGCGCTGGAGAGATGTCGAGATCTTGGCCCAATGGGAAATGTCGAGAGCA 187  
QY 202 CGAGCCCGAGTGAAGCACTACCCAGCCAGCAAGAGTAGCATTTCTGAGAGACTT 261  
Db 188 CGAGCCCGAGTGAAGCACTACCCAGCCAGCAAGAGTAGCATTTCTGAGAGACTT 247  
QY 262 CTTGATATAGGGCCATGATGCTCCCGAGGGGCTCTGGCCCCCGCAGAGAGGCAAG 321  
Db 248 CTTGATATAGGGCCATGATGCTCCCGAGGGGCTCTGGCCCCCGCAGAGAGGCAAG 307  
QY 322 TCCAGAGCTGGTGTCTGATGCTCCCTTCAATGACCTCAACCCGAGCTGTGAACAT 381  
Db 308 TCCAGAGCTGGTGTCTGATGCTCCCTTCAATGACCTCAACCCGAGCTGTGAACAT 367  
QY 382 GGGCTTTCGAGAGGATCTCGCAACATGCTGTGAAGCCGAGTGAACCTCACTGTCT 441  
Db 368 GGGCTTTCGAGAGGATCTCGCAACATGCTGTGAAGCCGAGTGAACCTCACTGTCT 427  
QY 442 CCTCTTCTGCTATGATGCTTGTGTGCTGCTCTCTCTGCTGTGCTGTGCTGTGCT 501  
Db 428 CCTCTTCTGCTATGATGCTTGTGTGCTGCTCTCTCTGCTGTGCTGTGCTGTGCT 487  
QY 502 GATGTCACCTGATGAGCGGTGACCTGAGAGGCTGATAGGGGTGCTTGTGTGAGAG 561  
Db 488 GATGTCACCTGATGAGCGGTGACCTGAGAGGCTGATAGGGGTGCTTGTGTGAGAG 547  
QY 562 GACCTGTGTGGCTTGTGTGAGAGCAAGCATTTTGAAGGGGATTTGTGTGCTTGG 621  
Db 548 GACCTGTGTGGCTTGTGTGAGAGCAAGCATTTTGAAGGGGATTTGTGTGCTTGG 607  
QY 622 AAGGATGATCAGAGAGGAGCAAGAGTGTGTTCCTCTTGTGTGAGCGTGAAGC 681

Db 608 AAGTATGATCGAGAGAGGAGACCAAGGTGTGTGTTCCCTTTGTGTAAAGCGAGGC 667  
Qy 682 AGAGGAGACGTTAGTCCAGCATTTCCAAAGTGTGGGTGGTCCGTGGTTCCGAGATA 741  
Db 668 AAGGAGAGCGTTAGTCCAGCATTTCCAAAGTGTGGGTGGTCCGTGGTTCCGAGATA 727  
Qy 742 CTTTATGTTGTGTGGGCTTGCATTAAGTGCACAAATACAGACAAAGAAAGCATGCC 801  
Db 728 CTTTATGTTGTGTGGGCTTGCATTAAGTGCACAAATACAGACAAAGAAAGCATGCC 787  
Qy 802 CTTCCCAATTTCTCTCAATCTTTTATGCGAGAAATCTCACTGGATGGCCAAATGTTTC 861  
Db 788 CTTCCCAATTTCTCTCAATCTTTTATGCGAGAAATCTCACTGGATGGCCAAATGTTTC 847  
Qy 862 CGATGCTGTGAGAGACATGCCAGCTCTCTGTGCTTATGAGAGAGAGACTTGGCTTAG 921  
Db 848 CGATGCTGTGAGAGACATGCCAGCTCTCTGTGCTTATGAGAGAGACTTGGCTTAG 907  
Qy 922 GCGAGGTGAGAAAAATTCAGACTTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 981  
Db 908 GCGAGGTGAGAAAAATTCAGACTTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 967  
Qy 982 ATTGGCTACTTTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1041  
Db 968 ATTGGCTACTTTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027  
Qy 1042 AATTAATAGATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1077  
Db 1028 AATTAATAGATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1063

RESULT 3  
US-09-809-391-58  
; Sequence 58, Application US/09809391  
; Publication No. US20030049618A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/809,391  
; CURRENT FILING DATE: 2001-03-16  
; Prior application data removed - consult PAMM or file wrapper  
; NUMBER OF SEQ ID NOS: 761  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 58  
; LENGTH: 1049  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-809-391-58

Query Match 66.4%; Score 792; DB 10; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
Qy 46 TGGCGCTGCGACACAGCATCTACTCAGCGGTGGTCACTCTGTGAACATCATGACTG 105  
Db 1 TGGCGCTGCGACACAGCATCTACTCAGCGGTGGTCACTCTGTGAACATCATGACTG 60  
Qy 106 CAAGCTCTCCCTCAATTTCTGTGTGAGCCCATCAGGAGCCCAAGCGCTGGAGAGATGT 165  
Db 61 CAAGCTCTCCCTCAATTTCTGTGTGAGCCCATCAGGAGCCCAAGCGCTGGAGAGATGT 120  
Qy 166 GCGGATCTTGGCCAAATGGGAAATGTGTGAGACACACACCCCGAGTGAAGACACTAC 225  
Db 121 GCGGATCTTGGCCAAATGGGAAATGTGTGAGACACACACCCCGAGTGAAGACACTAC 180  
Qy 226 CCAGGACCAAGAGAGATCTCTGTGACAGAGCTTCTTCAATAGAGGGCATGATGCTCC 285  
Db 181 CCAGGACCAAGAGAGATCTCTGTGACAGAGCTTCTTCAATAGAGGGCATGATGCTCC 240  
Qy 286 CCCAGGAGGATCTCTGTGAGCCCGCCAGACAGAGGAGGCTGGGCTGTGCTCACTC 345

Db 241 CCCAGGAGGATCTCTGTGAGCCCGCCAGACAGAGAGATGCCAGGCTGGGTCTGTCACTC 300  
Qy 346 CCCCTTCAATGACCTCAACCGGAGCTGTGTGAACATGAGGCTTTCCGAGTGGCATCTCG 405  
Db 301 CCCCTTCAATGACCTCAACCGGAGCTGTGTGAACATGAGGCTTTCCGAGTGGCATCTCG 360  
Qy 406 CAACATGCTGTGAGAGCCGCTGACCTTCACTCTGCTCTCTCTCTCTCATGATGCTTGG 465  
Db 361 CAACATGCTGTGAGAGCCGCTGACCTTCACTCTGCTCTCTCTCTCTCATGATGCTTGG 420  
Qy 466 TGTGTGTGAGGCTCTCTCTGAGGCTTGTGTGATCTGTGATCTGTGATCTGTGATCTGTG 525  
Db 421 TGTGTGTGAGGCTCTCTCTGAGGCTTGTGTGATCTGTGATCTGTGATCTGTGATCTGTG 480  
Qy 526 ACCTTGAAGGCTGTGATGAGGAGGCTTGTGTGATGAGGAGCTTGTGTGAGGCTTGTGTGAG 585  
Db 481 ACCTTGAAGGCTGTGATGAGGAGGCTTGTGTGATGAGGAGCTTGTGTGAGGCTTGTGTGAG 540  
Qy 586 AGCAGGATATTTGAGAGGAGATCTGTGTGAGGCTTGTGTGATGAGGAGGAGGAGGAGCA 645  
Db 541 AGCAGGATATTTGAGAGGAGATCTGTGTGAGGCTTGTGTGATGAGGAGGAGGAGGAGCA 600  
Qy 646 CAGGTGTGTGTTCCCTTGTGTGTTAAGCGTGAAGGAGGAGAGAGAGAGAGAGAGAGAG 704  
Db 601 CAGGTGTGTGTTCCCTTGTGTGTTAAGCGTGAAGGAGGAGAGAGAGAGAGAGAGAGAG 660  
Qy 705 TT-CCAAAGTGTGGGTGGTCCGTGTGTTCCGAGATACCTTTAAGTGTGTGTGAGGAGCTG 763  
Db 661 TTCCAAAGTGTGGGTGGTCCGTGTGTTCCGAGATACCTTTAAGTGTGTGTGAGGAGCTG 720  
Qy 764 CATTAAGTGCACAAATACAGAGCAAGAAAGCGATGCCCTTCCCATTTCTCTCATCTCT 823  
Db 721 CATTAAGTGCACAAATACAGAGCAAGAAAGCGATGCCCTTCCCATTTCTCTCATCTCT 780  
Qy 824 TTATGCGAGAGAGATCTCAGCTGTGATGCCAATGTTCCGATGCTGTGTGAGAGAGATGCC 883  
Db 781 TTATGCGAGAGAGATCTCAGCTGTGATGCCAATGTTCCGATGCTGTGTGAGAGAGATGCC 840  
Qy 884 GACGTCTCTCTGTGCTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 943  
Db 841 GACGTCTCTCTGTGCTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
Qy 944 CTTTATGAGACGTTTGTGTGTATGATGATATTTTATGCTACTTATGTTTATGAG 1003  
Db 901 CTTTATGAGACGTTTGTGTGTATGATGATATTTTATGCTACTTATGTTTATGAG 960  
Qy 1004 ACAAGTGTGAGCATCTATTTATGATGACCTTTCAATTAATAGATTAAAGTAAATA 1063  
Db 961 ACAAGTGTGAGCATCTATTTATGATGACCTTTCAATTAATAGATTAAAGTAAATA 1020  
Qy 1064 AAAAAAAAAAAAAA 1077  
Db 1021 AAAAAAAAAAAAAA 1034

RESULT 4  
US-09-882-171-58  
; Sequence 58, Application US/09882171  
; Publication No. US20030175858A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/882,171  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 09/809,391  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/149,476  
; PRIOR FILING DATE: 1998-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/04493  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: 60/040,162  
; PRIOR FILING DATE: 1997-03-07

Page 4

[illegible]



PRIOR APPLICATION NUMBER: 60/047,593  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,614  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/043,578  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/043,576  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/047,501  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/043,670  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/056,632  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,664  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,876  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,881  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,909  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,875  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,862  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,887  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,908  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/048,964  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/057,650  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: 60/056,884  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/057,669  
PRIOR FILING DATE: 1997-09-05

Query Match 66.4%; Score 792; DB 10; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
46 TCGGCGCTTGAGACAGACATCTACTAGGCTGGTCACTCTGTGAATCATCTGACTG 105  
1 TCGGCGCTTGAGACAGACATCTACTAGGCTGGTCACTCTGTGAATCATCTGACTG 60  
106 CAAAGCTTCCCTCAATTTCTGCTGAGCCATCAAGGACCCAGACGCGCTTGAGAGATGT 165  
61 CAAAGCTTCCCTCAATTTCTGCTGAGCCATCAAGGACCCAGACGCGCTTGAGAGATGT 120  
166 GCGGATCTTGCCCAATGGGAAATCTGTCAGAGAGAGACACCCCGAATGAGACCACTAC 225  
121 GCGGATCTTGCCCAATGGGAAATCTGTCAGAGAGAGACACCCCGAATGAGACCACTAC 180  
226 CCAAGCACAAGAGATGATCTTCTGACAGAGCTTCTTCAATAGGGGCCATGGTGTCTCC 285  
181 CCAAGCACAAGAGATGATCTTCTGACAGAGCTTCTTCAATAGGGGCCATGGTGTCTCC 240  
286 CCAAGGGGGTCTGCGCCCGCCAGACAGAGAGAGCTGCGAGCTGGGTGTCTCACTGTC 345  
241 CCAAGGGGGTCTGCGCCCGCCAGACAGAGAGAGCTGCGAGCTGGGTGTCTCACTGTC 300  
346 CCGCTTCAATACCTCAACCGGAGCTGGTGAATAGGGCTTTCGAGTGGCATCTCGG 405  
301 CCGCTTCAATACCTCAACCGGAGCTGGTGAATAGGGCTTTCGAGTGGCATCTCGG 360  
406 CAAAGATCTGAGAGCGGATGACCTTCATCTGCTCTTCTGCTCATGATGCTTGG 465  
361 CAAAGATCTGAGAGCGGATGACCTTCATCTGCTCTTCTGCTCATGATGCTTGG 420  
466 TGTTCGCGCTCTCTCTGCTTGGCTTGTCTACTGCTGCTCAACTGAGTCAAGCGGTG 525

421 TGTTCGCGCTCTCTCTGCTTGGCTTGTCTACTGCTGCTCAACTGAGTCAAGCGGTG 480  
526 ACCTGTAGAGGCTGATAGGGGTGGGTTTGTGAGAGGACTTGTGGGCTTGTGTGAG 585  
481 ACCTGTAGAGGCTGATAGGGGTGGGTTTGTGAGAGGACTTGTGGGCTTGTGTGAG 540  
586 AGCAGCATATTTGAGAGGATCTGCTGGTGGCTTGAAGGATATCATGAGAGGAGCCA 645  
541 AGCAGCATATTTGAGAGGATCTGCTGGTGGCTTGAAGGATATCATGAGAGGAGCCA 600  
646 CAGGTGTGTCTTCCCTTGTGTGAACGCTGAGGACAGAGGAGACGTTAGT-CCAGCAT 704  
601 CAGGTGTGTCTTCCCTTGTGTGAACGCTGAGGACAGAGGAGACGTTAGTCCAGCAT 660  
705 TT-CCAAAGTGTGGTGGTCCGTTGGTTCGCGAGATCTTTAGTGTGATGGGGCTG 763  
661 TTCCCAAGTGTGGTGGTCCGTTGGTTCGCGAGATCTTTAGTGTGATGGGGCTG 720  
764 CATTAGTGCAAAATCATAGACAAGAAAGCATGCGCTTCCAAATCTCTCAATCTT 823  
721 CATTAGTGCAAAATCATAGACAAGAAAGCATGCGCTTCCAAATCTCTCAATCTT 780  
824 TTATGCCGAGAAATCTGAGTGCATGATGCAATGTTCCAGTGCCTGTGAGAGACATGCC 883  
781 TTATGCCGAGAAATCTGAGTGCATGATGCAATGTTCCAGTGCCTGTGAGAGACATGCC 840  
884 GACGTCTCTCTGCTTGGTGGAGAGGAGACTTGGGCTTGGAGGAGTGAATAATTCAGA 943  
841 GACGTCTCTCTGCTTGGTGGAGAGGAGACTTGGGCTTGGAGGAGTGAATAATTCAGA 900  
944 CTTTGTAGCACTGTTTGTGTTTATGATGATATTTTATGCTACTTATGTTTGG 1003  
901 CTTTGTAGCACTGTTTGTGTTTATGATGATATTTTATGCTACTTATGTTTGG 960  
1004 ACAAGTGTAGTGGATCTATTTATGAGACCTTTGAATAAATGATTAAGTAAAAA 1063  
961 ACAAGTGTAGTGGATCTATTTATGATGACCTTTGAATAAATGATTAAGTAAAAA 1020

1064 AAAAAAAAAAAAAA 1077  
1021 AAAAAAAAAAAAAA 1034  
RESULT 5  
US-10-164-861-58  
Sequence 58, Application US/10164861  
Publication No. US20030225248A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/10/164,861  
CURRENT FILING DATE: 2002-06-10  
PRIOR FILING DATE: 1998-09-08  
PRIOR APPLICATION NUMBER: US/09/149,476  
PRIOR FILING DATE: 1998-03-06  
NUMBER OF SEQ ID NOS: 757  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 58  
LENGTH: 1049  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-164-861-58  
Query Match 66.4%; Score 792; DB 13; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
46 TCGGCGCTTGAGACAGACATCTACTAGGCTGGTCACTCTGTGAATCATCTGACTG 105  
1 TCGGCGCTTGAGACAGACATCTACTAGGCTGGTCACTCTGTGAATCATCTGACTG 60

106 CAAGCTCCCTCAATTTCTGGGAGAGCCCATCAGAGGACCCACAGCGCTGGAGATGCT 165  
61 CAAGCTCCCTCAATTTCTGGGAGAGCCCATCAGAGGACCCACAGCGCTGGAGATGCT 120  
166 GCGAGTCTTGGCCAAATGGGAAATCTGTCAGAGACGACGACCCCGAGTGAAGCACTAC 225  
121 GCGAGTCTTGGCCAAATGGGAAATCTGTCAGAGACGACGACCCCGAGTGAAGCACTAC 180  
226 CCAAGCCACCAAGAGTACATTTCTCCACAGAGCTTCTCAATGAGGAGCCATGCTCC 285  
181 CCAAGCCACCAAGAGTACATTTCTCCACAGAGCTTCTCAATGAGGAGCCATGCTCC 240  
286 CCAAGGAGGCTCTGGAGCCGCGCCAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 345  
241 CCAAGGAGGCTCTGGAGCCGCGCCAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
346 CCAAGTCTTCAATGAGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 405  
301 CCAAGTCTTCAATGAGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
406 CCAAGTCTTCAATGAGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 465  
361 CCAAGTCTTCAATGAGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
466 TGGTCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 525  
421 TGGTCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
526 ACCCTGAGGAGGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 585  
481 ACCCTGAGGAGGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
586 AGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 645  
541 AGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
646 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 704  
601 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
705 TT-CCAAAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 763  
661 TTCCAAAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
764 CATTAAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 823  
721 CATTAAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
824 TTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 883  
781 TTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
884 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 943  
841 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
944 CTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1003  
901 CTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
1004 ACAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1063  
961 ACAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
1064 AAAAAAAAAAAAAA 1077  
1021 AAAAAAAAAAAAAA 1034

GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Xue, Ailong J.  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and  
FILE REFERENCE: 803 1CNC2  
CURRENT APPLICATION NUMBER: US/10/302,172  
CURRENT FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/225,251  
PRIOR FILING DATE: 2002-08-20  
PRIOR APPLICATION NUMBER: PCT US02/05095  
PRIOR FILING DATE: 2002-03-05  
PRIOR APPLICATION NUMBER: US 09/799,451  
PRIOR FILING DATE: 2001-03-05  
NUMBER OF SEQ ID NOS: 950  
SOFTWARE: PL\_F1\_genes Version 2.0  
SEQ ID NO 94  
LENGTH: 1052  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (365)..(508)  
US-10-302-172-94

Query Match 56.1%; Score 669; DB 13; Length 1052;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 719; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

16 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 75  
1 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60  
76 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 135  
61 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120  
136 TCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 195  
121 TCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
196 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 255  
181 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
256 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 315  
241 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
316 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 375  
301 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
376 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 435  
361 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
436 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 495  
421 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
496 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 555  
481 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
556 TGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 615  
541 TGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
616 GCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 675

QY 676 TGAGGAGAGGAGAGAGCTTAGTCCAGCATTTCCAAAGTGTGGGTGGGTCCGTTGGTCC 735  
DB 661 TGAGGAGAGGAGAGAGCTTAGTCCAGCATTTCCAAAGTGTGGGTGGGTCCGTTGGTCC 720

RESULT 7  
US-10-076-747-65

/ Sequence 65, Application US/10076747  
/ Publication No. US20030180726A1

GENERAL INFORMATION:

/ APPLICANT: Salceda, Susana  
/ APPLICANT: Macina, Roberto  
/ APPLICANT: Hu, Ping  
/ APPLICANT: Recipon, Herve  
/ APPLICANT: Kaira, Kalpana  
/ APPLICANT: Cafferey, Robert  
/ APPLICANT: Sun, Yongming  
/ APPLICANT: Liu, Chenghua  
/ TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and  
/ FILE REFERENCE: DEX-0315  
/ CURRENT APPLICATION NUMBER: US/10/076,747  
/ CURRENT FILING DATE: 2002-02-13  
/ PRIOR APPLICATION NUMBER: 60/268,290  
/ PRIOR FILING DATE: 2001-02-13  
/ PRIOR APPLICATION NUMBER: 60/268,834  
/ PRIOR FILING DATE: 2001-02-15  
/ NUMBER OF SEQ ID NOS: 129  
/ SOFTWARE: PatentIn version 3.1  
/ SEQ ID NO 65  
/ LENGTH: 365  
/ TYPE: DNA  
/ ORGANISM: Homo sapien  
US-10-076-747-65

Query Match  
Best Local Similarity 100.0%; Score 326; DB 15; Length 365;  
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 GACGACCCCGAGTGAAGACCTACCCAGCCAGCAAGAGTGAATCTCTGACAGAGC 259  
DB 40 GACGACCCCGAGTGAAGACCTACCCAGCCAGCAAGAGTGAATCTCTGACAGAGC 99  
QY 260 TTCTTAATAGGGGGCATGTGTCTCCCGAGGGGTCTGGCCCGGAGAGAGAGCA 319  
DB 100 TTCTTAATAGGGGGCATGTGTCTCCCGAGGGGTCTGGCCCGGAGAGAGAGCA 159  
QY 320 GGTGCAAGCTGGGTGTGTCTGATGCTCCCTTCATGACTCAACCGAGAGTGTGAAC 379  
DB 160 GGTGCAAGCTGGGTGTGTCTGATGCTCCCTTCATGACTCAACCGAGAGTGTGAAC 219  
QY 380 ATGGGCTTTCCGAGAGGCAATCTCCGGAACATGCTGTGAAGCCGCTGATCTTCATCTG 439  
DB 220 ATGGGCTTTCCGAGAGGCAATCTCCGGAACATGCTGTGAAGCCGCTGATCTTCATCTG 279  
QY 440 CTCCTCTTCTGCTCATGATGCTGTGCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTG 499  
DB 280 CTCCTCTTCTGCTCATGATGCTGTGCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTG 339  
QY 500 CTGGTGTCCCACTGAGTCAAGCGGTG 525  
DB 340 CTGGTGTCCCACTGAGTCAAGCGGTG 365

RESULT 8  
US-10-076-747-64

/ Sequence 64, Application US/10076747  
/ Publication No. US20030180726A1

GENERAL INFORMATION:

/ APPLICANT: Salceda, Susana  
/ APPLICANT: Macina, Roberto  
/ APPLICANT: Hu, Ping  
/ APPLICANT: Recipon, Herve

/ APPLICANT: Kaira, Kalpana  
/ APPLICANT: Cafferey, Robert  
/ APPLICANT: Sun, Yongming  
/ APPLICANT: Liu, Chenghua  
/ TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and P  
/ FILE REFERENCE: DEX-0315  
/ CURRENT APPLICATION NUMBER: US/10/076,747  
/ CURRENT FILING DATE: 2002-02-13  
/ PRIOR APPLICATION NUMBER: 60/268,290  
/ PRIOR FILING DATE: 2001-02-13  
/ PRIOR APPLICATION NUMBER: 60/268,834  
/ PRIOR FILING DATE: 2001-02-15  
/ NUMBER OF SEQ ID NOS: 129  
/ SOFTWARE: PatentIn version 3.1  
/ SEQ ID NO 64  
/ LENGTH: 612  
/ TYPE: DNA  
/ ORGANISM: Homo sapien  
US-10-076-747-64

Query Match  
Best Local Similarity 100.0%; Score 195; DB 15; Length 612;  
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 828 GCCGAGAGATCTAGCTGATGCGCAACATGTCGATGCTGTGGAAGATGCCGAGC 887  
DB 261 GCCGAGAGATCTAGCTGATGCGCAACATGTCGATGCTGTGGAAGATGCCGAGC 320  
QY 888 TCTCCTGCTTGAAGAGAGAGAGCTTGGGCTTGAAGAGAGAGAGAGAGAGAGAGAG 947  
DB 321 TCTCCTGCTTGAAGAGAGAGAGAGCTTGGGCTTGAAGAGAGAGAGAGAGAGAGAG 380  
QY 948 TTAGACATGTTTGTGTTTAAATGATATTTTATGCTACTTATGTTTGAAGAA 1007  
DB 381 TTAGACATGTTTGTGTTTAAATGATATTTTATGCTACTTATGTTTGAAGAA 440  
QY 1008 GTGTAGTGCATTC 1022  
DB 441 GTGTAGTGCATTC 455

RESULT 9  
US-10-001-885-30

/ Sequence 30, Application US/10001885  
/ Publication No. US20040058319A1

GENERAL INFORMATION:

/ APPLICANT: Salceda, Susana  
/ APPLICANT: Macina, Roberto  
/ APPLICANT: Recipon, Herve  
/ APPLICANT: Cafferey, Robert  
/ APPLICANT: Sun, Yongming  
/ APPLICANT: Liu, Chenghua  
/ TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro  
/ FILE REFERENCE: DEX-0279  
/ CURRENT APPLICATION NUMBER: US/10/001,885  
/ CURRENT FILING DATE: 2001-11-20  
/ PRIOR APPLICATION NUMBER: 60/252,061  
/ PRIOR FILING DATE: 2000-11-20  
/ PRIOR APPLICATION NUMBER: 60/253,257  
/ PRIOR FILING DATE: 2000-11-27  
/ NUMBER OF SEQ ID NOS: 167  
/ SOFTWARE: PatentIn version 3.1  
/ SEQ ID NO 30  
/ LENGTH: 643  
/ TYPE: DNA  
/ ORGANISM: Homo sapien  
US-10-001-885-30

Query Match  
Best Local Similarity 100.0%; Score 125; DB 13; Length 643;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 828 GCCGAGAGATCTAGCTGATGCGCAACATGTCGATGCTGTGGAAGATGCCGAGC 887



OTHER INFORMATION: MAP TO AC004514.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
US-09-864-761-6261

Query Match 3.2%; Score 38; DB 9; Length 473;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 TAGATTTAAGTAAAAA  
DB 259 TAGATTTAAGTAAAAA  
1085  
296

RESULT 12  
US-10-027-632-209023

Sequence 209023, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 209023  
LENGTH: 998  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-209023

Query Match 3.2%; Score 38; DB 13; Length 998;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 TAGATTTAAGTAAAAA  
DB 307 TAGATTTAAGTAAAAA  
1085  
344

RESULT 13  
US-10-027-632-209023

Sequence 209023, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 209023  
LENGTH: 998  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-209023

Query Match 3.2%; Score 38; DB 16; Length 998;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 TAGATTTAAGTAAAAA  
DB 307 TAGATTTAAGTAAAAA  
1085  
344

RESULT 14  
US-10-424-599-125925/C  
Sequence 125925, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 125925  
LENGTH: 616  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_84719C.1  
US-10-424-599-125925

Query Match 2.9%; Score 34; DB 13; Length 616;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TTTAAGTAAAAA  
DB 65 TTTAAGTAAAAA  
1085  
32

RESULT 15  
US-10-678-521-24  
Sequence 24, Application US/10678521  
Publication No. US20040081998A1  
GENERAL INFORMATION:  
APPLICANT: Silver, Gary M.  
APPLICANT: Brandt, Kevin S.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: NOVEL CARBOXYL ESTERASE NUCLEIC ACID MOLECULES, PROTEINS AND USES  
FILE REFERENCE: FC-1-CI-FUS

Tue Aug 24 09:48:32 2004

us-10-001-885-31.011.rmpb

Page 10

CURRENT APPLICATION NUMBER: US/10/678,521  
CURRENT FILING DATE: 2003-10-02  
PRIOR APPLICATION NUMBER: US/09/403,942F  
PRIOR FILING DATE: 2000-05-02  
PRIOR APPLICATION NUMBER: PCT/US97/20598  
PRIOR FILING DATE: 1997-11-10  
PRIOR APPLICATION NUMBER: 08/747,221  
PRIOR FILING DATE: 1996-11-12  
NUMBER OF SEQ ID NOS: 76  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 24  
LENGTH: 2836  
TYPE: DNA  
ORGANISM: Ctenocephalides felis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (99)..(1889)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (2278)..(2278)  
OTHER INFORMATION: n = unknown  
US-10-678-521-24

Query Match 2.98; Score 34; DB 17; Length 2836;  
Best Local Similarity 100.0%; Pred.No. 1.2e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1052 TTTAAGTAAAAA 1085  
Db 2778 TTTAAGTAAAAA 2811

Search completed: August 22, 2004, 08:31:53  
Job time : 599 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:04 ; Search time 54 Seconds

(without alignments)  
633.116 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632  
Sequence: 1 MVRILANGRIYQDDPRVRT.....GYRGILLVLYVLSHSQR 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 632   | 100.0       | 121    | 5     | AB079328 Human ova  |
| 2          | 632   | 100.0       | 121    | 6     | AB061063 Human ova  |
| 3          | 568.5 | 90.0        | 120    | 7     | AD060341 Rat Prote  |
| 4          | 568.5 | 90.0        | 120    | 7     | AD060339 Rat Prote  |
| 5          | 387   | 61.2        | 111    | 5     | AB035533 Human nov  |
| 6          | 387   | 61.2        | 111    | 6     | AB034727 Fragment   |
| 7          | 245   | 38.8        | 48     | 2     | AA074777 Human sec  |
| 8          | 245   | 38.8        | 48     | 5     | AB068995 Human pol  |
| 9          | 245   | 38.8        | 48     | 5     | AB052227 Human nov  |
| 10         | 245   | 38.8        | 48     | 6     | AB034421 Region of  |
| 11         | 184   | 29.1        | 114    | 4     | AB030633 Novel hum  |
| 12         | 110.5 | 17.5        | 132    | 6     | AB084669 Human SEC  |
| 13         | 110.5 | 17.5        | 132    | 7     | AD080804 Novel pro  |
| 14         | 80    | 12.7        | 92     | 4     | AA077595 Human col  |
| 15         | 80    | 12.7        | 92     | 4     | AA042404 Human ORF  |
| 16         | 78.5  | 12.4        | 325    | 2     | AA041419 Paxillin   |
| 17         | 78.5  | 12.3        | 545    | 2     | AA027493 p34 from   |
| 18         | 77.5  | 12.3        | 1568   | 4     | AB0303109 Novel hum |
| 19         | 75.5  | 11.9        | 622    | 4     | AA068560 Human nov  |
| 20         | 75.5  | 11.9        | 622    | 4     | AB042888 Novel hum  |
| 21         | 75.5  | 11.9        | 622    | 7     | AD033068 Human nov  |
| 22         | 75    | 11.9        | 417    | 3     | AA012139 Hydropob   |
| 23         | 75    | 11.9        | 718    | 3     | AA071059 Human mem  |
| 24         | 75    | 11.9        | 718    | 5     | AB052104 Homo sapi  |
| 25         | 75    | 11.9        | 735    | 4     | AB035546 Human pro  |

#### ALIGNMENTS

|    |      |      |       |   |                     |
|----|------|------|-------|---|---------------------|
| 26 | 73.5 | 11.6 | 442   | 4 | AB0909021 Novel hum |
| 27 | 73   | 11.6 | 763   | 3 | AA042392 Arabidops  |
| 28 | 73   | 11.6 | 793   | 3 | AA042391 Arabidops  |
| 29 | 72.5 | 11.5 | 315   | 4 | AB061702 Drosophi   |
| 30 | 72   | 11.3 | 310   | 2 | AA023996 B-HDAg-T   |
| 31 | 71.5 | 11.3 | 158   | 3 | AA022606 Zea mays   |
| 32 | 71.5 | 11.3 | 19938 | 6 | AB076680 Streptomy  |
| 33 | 71   | 11.2 | 124   | 4 | AA001631 Human pol  |
| 34 | 71   | 11.2 | 345   | 3 | AA093600 Protein e  |
| 35 | 71   | 11.2 | 349   | 4 | AB04137 Novel hum   |
| 36 | 71   | 11.2 | 689   | 2 | AA072025 HSV-2 str  |
| 37 | 71   | 11.2 | 1030  | 4 | AB022699 Novel hum  |
| 38 | 71   | 11.2 | 1697  | 7 | AD037542 Human nuc  |
| 39 | 70.5 | 11.2 | 226   | 3 | AA042794 Human CRF  |
| 40 | 70.5 | 11.2 | 915   | 5 | AB063120 FLO11 gen  |
| 41 | 70   | 11.1 | 213   | 6 | AB073050 Amino aci  |
| 42 | 70   | 11.1 | 266   | 4 | AB018528 Novel hum  |
| 43 | 70   | 11.1 | 379   | 7 | AD085283 Rat foocce |
| 44 | 70   | 11.1 | 387   | 4 | AB011608 Novel hum  |
| 45 | 70   | 11.1 | 1128  | 4 | AB058911 Drosophi   |

RESULT 1  
AB079328  
ID AB079328 standard; protein, 121 AA.  
XX  
AC AB079328;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE Human ovary specific protein SEQ ID NO:125.  
XX  
KW Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;  
KW ovary specific gene; OSN; ovarian cancer; immune response; metastasis.  
XX  
OS Homo sapiens.  
XX  
PN WO200240535-A2.  
XX  
PD 23-MAY-2002.  
XX  
PF 20-NOV-2001; 2001WO-US045011.  
XX  
PR 20-NOV-2000; 2000US-0252061P.  
PR 27-NOV-2000; 2000US-0253257P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Salceda S, Macina RA, Recipon H, Caffertkey R, Sun Y, Liu C;  
XX  
DR WPI, 2002-471617/50.  
XX  
PT New ovary specific genes and proteins, useful as a vaccine for treating  
PT patients with ovarian cancer, or for diagnosing and monitoring the  
PT presence and metastases of ovarian cancer in a patient.  
XX  
PS Claim 11, Page 241; 260pp; English.  
XX  
XX AB087790 to AB087882 represent human ovary specific nucleic acid (OSNA)  
XX sequences, and AB079297 to AB079370 represent ovary specific protein  
XX (OSP) sequences from the present invention. OSNA and OSP sequences have  
XX cytostatic activity, and can be used in vaccine production and gene  
XX therapy. An antibody, that specifically binds to an OSP can be used for  
XX treating a patient with ovarian cancer, particularly for inducing an  
XX immune response against the ovarian cancer cell expressing the OSNA or  
XX OSP. The OSNAs and OSPs can also be used for diagnosing and monitoring  
XX the presence and metastases of ovarian cancer in a patient  
XX Sequence 121 AA;  
XX



Query Match 100.0%; Score 632; DB 5; Length 121;  
Best Local Similarity 100.0%; Pred. No. 8.9e-62;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGELIVODDDPRVRRITTOPPGRSIPROSFENRGNHGAAPGCGPPOQAGARLGA 60  
DB 1 MVRILANGELIVODDDPRVRRITTOPPGRSIPROSFENRGNHGAAPGCGPPOQAGARLGA 60  
QY 61 QSPFNDLRQLVNNGFPOWHLGNHAEVPTSIILLFLMMLGVKGLLLVGLVYVSHLSQ 120  
DB 61 QSPFNDLRQLVNNGFPOWHLGNHAEVPTSIILLFLMMLGVKGLLLVGLVYVSHLSQ 120  
QY 121 R 121  
DB 121 R 121

## RESULT 2

ABU61063  
ID ABU61063 standard; protein; 121 AA.

XX ABU61063;

DT 08-MAY-2003 (first entry)

DE Human ovarian specific protein DEX0310\_122.

XX Human; ovarian specific protein; OSP; ovarian cancer;  
KW non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;  
KM gynaecological.

OS Homo sapiens.

PN W0200292785-A2.

PD 21-NOV-2002.

PF 13-FEB-2002; 2002MO-US022271.

PR 13-FEB-2001; 2001US-0268290P.

PR 15-FEB-2001; 2001US-0268834P.

PA (DIAD-) DIADEXUS INC.

PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;  
PI Sun Y, Liu C;

XX WPI; 2003-120677/11.

DR N-PSDB; ABX92251.

PT New isolated OSNA nucleic acid and encoded polypeptide, useful for  
PT identifying, diagnosing, monitoring, staging, imaging and treating  
PT ovarian cancer and non-cancerous diseases in ovarian tissues.

PS Claim 11; Page 220-221; 224pp; English.

XX The invention relates to a new isolated nucleic acid termed ovarian

CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that

CC encodes any of 53 fully defined protein sequences appearing as ABU61018-

CC ABU61070 (termed ovarian specific proteins, OSP); (b) any of 76 fully

CC defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a

CC molecule of (a) or (b). Also included are a method for determining the

CC presence of an OSNA, a host cell comprising the vector, an isolated OSP

CC comprising an OSNA, an anti-OSP antibody or fragment, a method for determining  
CC the presence of an OSNA, a method for identifying a sample and a vaccine  
CC comprising an OSP or OSNA. The methods and compositions of the present  
CC invention are useful for identifying, diagnosing, monitoring, staging,  
CC imaging and treating ovarian cancer and non-cancerous disease in ovary  
CC tissue. The present sequence is an OSP of the invention

XX Sequence 121 AA;

Query Match 100.0%; Score 632; DB 6; Length 121;  
Best Local Similarity 100.0%; Pred. No. 8.9e-62;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGELIVODDDPRVRRITTOPPGRSIPROSFENRGNHGAAPGCGPPOQAGARLGA 60  
DB 1 MVRILANGELIVODDDPRVRRITTOPPGRSIPROSFENRGNHGAAPGCGPPOQAGARLGA 60  
QY 61 QSPFNDLRQLVNNGFPOWHLGNHAEVPTSIILLFLMMLGVKGLLLVGLVYVSHLSQ 120  
DB 61 QSPFNDLRQLVNNGFPOWHLGNHAEVPTSIILLFLMMLGVKGLLLVGLVYVSHLSQ 120  
QY 121 R 121  
DB 121 R 121

## RESULT 3

ADE60341  
ID ADE60341 standard; protein; 120 AA.

XX ADE60341;

DT 29-JAN-2004 (first entry)

DE Rat Protein BAB25613, SEQ ID NO 6247.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN W02003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002MO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GENO) GEN HOSPITAL CORP.

PI (FARB) BAYER AG.

XX Woolf C, D'Urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.

DR GENBANK; BAB25613.

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PT

PS Claim 1; Page: 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a

CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 120 AA:

Query Match 90.0%; Score 568.5; DB 7; Length 120;

Best Local Similarity 92.6%; Pred. No. 9.1e-55;

Matches 112; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MRLIANGELIVDDPRVETTTQPPRGSIHQSFENRGGAPPGGPGPQQOAGARLGAA 60

DB 1 MRLIANGELIVDDPRVETTTQ-HRSSSSQCGFENRGGAPPGGPGPQQOAGARLGAA 59

QY 61 QSPFDLNRQLVNMGGPQWHLGNHVEPVTSILLFLMLMGVGLLVGLVIVSHLSQ 120

DB 60 QSPFDLNRQLVNMGGPQWHLGNHVEPVTSILLFLMLMGVGLLVGLVIVSHLSQ 119

QY 121 R 121

DB 120 R 120

RESULT 4  
ADE60339 standard; protein; 120 AA.

AC ADE60339;

DT 29-JAN-2004 (first entry)

XX Rat Protein BAB5613, SEQ ID NO 6245.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

OS MO2003016475-A2.

PN 27-FEB-2003.

PF 14-AUG-2002; 2002WC-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

XX (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'Urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; BAB5613.

XX New composition comprising two or more isolated polypeptides, useful for

XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017P; English.

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotide or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 120 AA:

Query Match 90.0%; Score 568.5; DB 7; Length 120;

Best Local Similarity 92.6%; Pred. No. 9.1e-55;

Matches 112; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MRLIANGELIVDDPRVETTTQPPRGSIHQSFENRGGAPPGGPGPQQOAGARLGAA 60

DB 1 MRLIANGELIVDDPRVETTTQ-HRSSSSQCGFENRGGAPPGGPGPQQOAGARLGAA 59

QY 61 QSPFDLNRQLVNMGGPQWHLGNHVEPVTSILLFLMLMGVGLLVGLVIVSHLSQ 120

DB 60 QSPFDLNRQLVNMGGPQWHLGNHVEPVTSILLFLMLMGVGLLVGLVIVSHLSQ 119

QY 121 R 121

DB 120 R 120

RESULT 5  
ABG95533 standard; protein; 111 AA.

AC ABG95533;

DT 15-JAN-2003 (first entry)

XX Human novel secreted protein gene 48 polypeptide #1.

XX Human; secreted protein; autoimmune disease; chemotaxis;

KM rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;

KM liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;

KM cerebrovascular disorder; cerebral ischaemia; angiogenesis; sunburn;

KM nervous system disorders; Alzheimer's disease; infection;

KM ocular disorder; corneal infection; wound healing; tissue regeneration;

KM epithelial cell proliferation; organ transplantation; food additive;

XX preservative; nutritional.

OS Homo sapiens.

PN US6420526-B1.

PD 16-JUL-2002.

PF 08-SEP-1998; 98US-00149476.

PR 07-MAR-1997; 97US-0038621P.

PR 07-MAR-1997; 97US-0040161P.

PR 07-MAR-1997; 97US-0040162P.

PR 07-MAR-1997; 97US-0040163P.



Db 99 GSEFNDLNRLVNV 111

RESULT 6  
ABO34727  
ID ABO34727 standard; protein; 111 AA.  
XX ABO34727;  
AC  
XX 22-SEP-2003 (first entry)  
DT  
XX Fragment #45 of a human secreted protein.  
XX  
KW Human; secreted protein; hyperproliferative disorder; leukaemia;  
KW breast cancer; wound; reproductive disorder; blood-related disorder;  
KW haemophilia; thrombocytopaenia; immunodeficiency; chronic hypoplasia;  
KW Wiskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;  
KW graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;  
KW viral infection; bacterial infection; fungal infection; AIDS; sepsis;  
KW renal disorder; kidney failure; cardiovascular disorder; cytostatic;  
KW angina pectoris; cerebral ischaemia; congenital heart defect;  
KW respiratory disorder; neurological disorder; Alzheimer's disease;  
KW Parkinson's disease; inflammation; Crohn's disease; vulnery;  
KW immunosuppressive; antibacterial; haemostatic; thrombolytic;  
KW anticoagulant; neuroprotective; thymolytic; antiallergic;  
KW antidiabetic; virucide; fungicide; anti-HIV; neurotrophic; antiangiinal;  
KW cerebroprotective; cardiant; nootropic; antiparkinsonian;  
KW antiinflammatory.  
XX Homo sapiens.  
XX  
XX US2003049618-A1.  
XX  
XX 13-MAR-2003.  
XX  
XX 16-MAR-2001; 2001US-00809391.  
XX  
PR 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040335P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043659P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0047482P.  
PR 23-MAY-1997; 97US-0047500P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.  
PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047504P.  
PR 23-MAY-1997; 97US-0047581P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.

PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 06-JUN-1997; 97US-0048974P.  
PR 13-JUN-1997; 97US-0049610P.  
PR 08-JUL-1997; 97US-0051926P.  
PR 16-JUL-1997; 97US-0052872P.  
PR 16-AUG-1997; 97US-0052724P.  
PR 22-AUG-1997; 97US-005630P.  
PR 22-AUG-1997; 97US-0056631P.  
PR 22-AUG-1997; 97US-0056632P.  
PR 22-AUG-1997; 97US-0056635P.  
PR 22-AUG-1997; 97US-0056637P.  
PR 22-AUG-1997; 97US-0056662P.  
PR 22-AUG-1997; 97US-0056664P.  
PR 22-AUG-1997; 97US-0056845P.  
PR 22-AUG-1997; 97US-0056862P.  
PR 22-AUG-1997; 97US-0056864P.  
PR 22-AUG-1997; 97US-0056872P.  
PR 22-AUG-1997; 97US-0056874P.  
PR 22-AUG-1997; 97US-0056875P.  
PR 22-AUG-1997; 97US-0056876P.  
PR 22-AUG-1997; 97US-0056877P.  
PR 22-AUG-1997; 97US-0056878P.  
PR 22-AUG-1997; 97US-0056879P.  
PR 22-AUG-1997; 97US-0056880P.  
PR 22-AUG-1997; 97US-0056881P.  
PR 22-AUG-1997; 97US-0056882P.  
PR 22-AUG-1997; 97US-0056884P.  
PR 22-AUG-1997; 97US-0056886P.  
PR 22-AUG-1997; 97US-0056887P.  
PR 22-AUG-1997; 97US-0056888P.  
PR 22-AUG-1997; 97US-0056889P.  
PR 22-AUG-1997; 97US-0056892P.  
PR 22-AUG-1997; 97US-0056893P.  
PR 22-AUG-1997; 97US-0056894P.  
PR 22-AUG-1997; 97US-0056903P.  
PR 22-AUG-1997; 97US-0056908P.  
PR 22-AUG-1997; 97US-0056909P.  
PR 22-AUG-1997; 97US-0056910P.  
PR 22-AUG-1997; 97US-0056911P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057659P.  
PR 12-SEP-1997; 97US-0057761P.  
PR 12-SEP-1997; 97US-0058785P.  
PR 09-OCT-1997; 97US-0061660P.  
PR 06-MAR-1998; 98WO-US004493.  
PR 08-SEP-1998; 98US-0014947G.  
PR 17-MAR-2000; 2000US-0190068P.  
XX  
XX (RUBE/) RUBEN S M.  
PA (ROSE/) ROSEN C A.  
PA (SOPP/) SOPPET D R.  
PA (CART/) CARTER K C.  
PA (BEDN/) BEDNARIK D P.  
PA (ENDR/) ENDRESS G A.

PA (YUGG/) YU G.  
 PA (NIJJ/) NI J.  
 PA (FENG/) FENG P.  
 PA (YOUN/) YOUNG P. E.  
 PA (GREE/) GREENE J. M.  
 PA (FERE/) FERRIE A. M.  
 PA (DUAN/) DUAN D. R.  
 PA (HUJ/) HU J.  
 PA (FLOP/) FLORENCE K. A.  
 PA (OLSE/) OLSEN H. S.  
 PA (FISC/) FISCHER C. L.  
 PA (EBEN/) EBNER R.  
 PA (BREW/) BREWER L. A.  
 PA (MOOR/) MOORE P. A.  
 PA (SHIY/) SHI Y.  
 PA (LAFI/) LAFLEUR D. W.  
 PA (LIYY/) LI Y.  
 PA (ZENG/) ZENG Z.  
 PA (KJAW/) KJAW H.  
 XX  
 XX Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP,  
 PI Endress GA, Yu G, Ni J, Feng P, Young PE, Greene UM, Ferrie AM;  
 PI Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R;  
 PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
 XX WPI; 2003-521800/49.  
 XX  
 XX New genes and its encoded prostate cancer antigen proteins, useful for  
 PT preventing, treating, ameliorating or diagnosing e.g. prostate cancers,  
 PT lymph hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral  
 PT ischemia.  
 XX  
 XX Claim 3; Page 29; 260pp; English.  
 XX  
 CC The present invention relates to the isolation of novel human secreted  
 CC proteins and the polynucleotide sequences encoding them. The invention  
 CC also discloses vectors, host cells, antibodies, and recombinant methods  
 CC for producing human secreted proteins. The polypeptide and polynucleotide  
 CC sequences for the secreted proteins are useful for preventing, treating,  
 CC ameliorating or diagnosing medical conditions such as hyperproliferative  
 CC disorders (e.g. leukemia or breast cancers), wounds, reproductive  
 CC disorders, blood-related disorders (e.g. haemophilia or  
 CC thrombocytopenia), immunodeficiencies (e.g. Wiskott-Aldrich syndrome or  
 CC lymphic hypoplasia), autoimmune disorders (e.g. graft-versus-host disease,  
 CC multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g. asthma),  
 CC vital or bacterial or fungal infections (e.g. AIDS or sepsis), renal  
 CC disorders (e.g. kidney failure), cardiovascular disorders (e.g. angina  
 CC pectoris, cerebral ischaemia or congenital heart defects), respiratory  
 CC disorders, neurological disorders (e.g. Alzheimer's disease or  
 CC Parkinson's disease), and inflammations (e.g. Crohn's disease). The  
 CC polynucleotide or polypeptide may also be used as vaccine adjuvants.  
 CC AB034374-AB034815 represent human secreted proteins or their fragments.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO web site at: seqdata.uspto.gov/patidentry.html  
 XX  
 XX Sequence 111 AA;

Query Match 61.2%; Score 387; DB 6; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 9,5e-35;  
 Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGELVODDDPRVTTTTPPRGSIROSFFNRGHCAPGCGPRQOQAGALGAA 60  
 DB 39 MVRILANGELVODDDPRVTTTTPPRGSIROSFFNRGHCAPGCGPRQOQAGALGAA 98  
 QY 61 QSPENDINROLVN 73  
 DB 99 QSPENDINROLVN 111

RESULT 7  
 AAW74777

ID AAW74777 standard; protein; 48 AA.  
 XX  
 AC AAW74777;  
 XX  
 DT 19-JAN-1999 (first entry)  
 XX  
 DE Human secreted protein encoded by gene 48 clone HFCAT74.  
 XX  
 XX Human; secreted protein; testis; tumour; foetal brain tissue;  
 KM fusion protein; cancer; central nervous system; seizure; diagnosis;  
 KM neurodegenerative disease.  
 XX  
 OS Homo sapiens.  
 XX  
 XX MO9839448-A2.  
 XX  
 PD 11-SEP-1998.  
 XX  
 PF 06-MAR-1998; 98MO-US004493.  
 XX  
 XX 07-MAR-1997; 97US-0038621P.  
 PR 07-MAR-1997; 97US-0040161P.  
 PR 07-MAR-1997; 97US-0040162P.  
 PR 07-MAR-1997; 97US-0040163P.  
 PR 07-MAR-1997; 97US-0040333P.  
 PR 07-MAR-1997; 97US-0040334P.  
 PR 07-MAR-1997; 97US-0040335P.  
 PR 07-MAR-1997; 97US-0040626P.  
 PR 11-APR-1997; 97US-0043311P.  
 PR 11-APR-1997; 97US-0043312P.  
 PR 11-APR-1997; 97US-0043313P.  
 PR 11-APR-1997; 97US-0043314P.  
 PR 11-APR-1997; 97US-0043315P.  
 PR 11-APR-1997; 97US-0043568P.  
 PR 11-APR-1997; 97US-0043569P.  
 PR 11-APR-1997; 97US-0043576P.  
 PR 11-APR-1997; 97US-0043578P.  
 PR 11-APR-1997; 97US-0043580P.  
 PR 11-APR-1997; 97US-0043669P.  
 PR 11-APR-1997; 97US-0043670P.  
 PR 11-APR-1997; 97US-0043671P.  
 PR 11-APR-1997; 97US-0043672P.  
 PR 11-APR-1997; 97US-0043674P.  
 PR 23-MAY-1997; 97US-0047492P.  
 PR 23-MAY-1997; 97US-0047500P.  
 PR 23-MAY-1997; 97US-0047501P.  
 PR 23-MAY-1997; 97US-0047502P.  
 PR 23-MAY-1997; 97US-0047503P.  
 PR 23-MAY-1997; 97US-0047504P.  
 PR 23-MAY-1997; 97US-0047505P.  
 PR 23-MAY-1997; 97US-0047506P.  
 PR 23-MAY-1997; 97US-0047507P.  
 PR 23-MAY-1997; 97US-0047508P.  
 PR 23-MAY-1997; 97US-0047509P.  
 PR 23-MAY-1997; 97US-0047510P.  
 PR 23-MAY-1997; 97US-0047511P.  
 PR 23-MAY-1997; 97US-0047512P.  
 PR 23-MAY-1997; 97US-0047513P.  
 PR 23-MAY-1997; 97US-0047514P.  
 PR 23-MAY-1997; 97US-0047515P.  
 PR 23-MAY-1997; 97US-0047516P.

PR 23-MAY-1997; 97US-0047618P.  
 PR 23-MAY-1997; 97US-0047632P.  
 PR 23-MAY-1997; 97US-0047633P.  
 PR 06-JUN-1997; 97US-0048964P.  
 PR 06-JUN-1997; 97US-0048974P.  
 PR 13-JUN-1997; 97US-0049610P.  
 PR 08-JUL-1997; 97US-0051926P.  
 PR 16-JUL-1997; 97US-0052874P.  
 PR 18-AUG-1997; 97US-0055724P.  
 PR 22-AUG-1997; 97US-0056630P.  
 PR 22-AUG-1997; 97US-0056631P.  
 PR 22-AUG-1997; 97US-0056632P.  
 PR 22-AUG-1997; 97US-0056636P.  
 PR 22-AUG-1997; 97US-0056637P.  
 PR 22-AUG-1997; 97US-0056662P.  
 PR 22-AUG-1997; 97US-0056664P.  
 PR 22-AUG-1997; 97US-0056684P.  
 PR 22-AUG-1997; 97US-0056686P.  
 PR 22-AUG-1997; 97US-0056872P.  
 PR 22-AUG-1997; 97US-0056874P.  
 PR 22-AUG-1997; 97US-0056875P.  
 PR 22-AUG-1997; 97US-0056876P.  
 PR 22-AUG-1997; 97US-0056877P.  
 PR 22-AUG-1997; 97US-0056878P.  
 PR 22-AUG-1997; 97US-0056879P.  
 PR 22-AUG-1997; 97US-0056880P.  
 PR 22-AUG-1997; 97US-0056881P.  
 PR 22-AUG-1997; 97US-0056882P.  
 PR 22-AUG-1997; 97US-0056884P.  
 PR 22-AUG-1997; 97US-0056886P.  
 PR 22-AUG-1997; 97US-0056887P.  
 PR 22-AUG-1997; 97US-0056888P.  
 PR 22-AUG-1997; 97US-0056889P.  
 PR 22-AUG-1997; 97US-0056892P.  
 PR 22-AUG-1997; 97US-0056893P.  
 PR 22-AUG-1997; 97US-0056894P.  
 PR 22-AUG-1997; 97US-0056903P.  
 PR 22-AUG-1997; 97US-0056908P.  
 PR 22-AUG-1997; 97US-0056909P.  
 PR 22-AUG-1997; 97US-0056910P.  
 PR 22-AUG-1997; 97US-0056911P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057659P.  
 PR 05-SEP-1997; 97US-0057761P.  
 PR 12-SEP-1997; 97US-0058785P.  
 PR 02-OCT-1997; 97US-0061060P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC,  
 PI Bednarik DP, Endress GA, Yu G, Ni J, Feng P, Young PS, Greene JM,  
 PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA,  
 PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
 XX WPI, 1998-506364/43.  
 DR N-PSDB; AAVS9558.  
 DR  
 XX  
 PT New isolated human genes and the secreted polypeptide(s) they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders.  
 XX  
 PS Claim 1; Page 567; 721pp; English.  
 XX  
 CC This sequence represents a secreted human protein encoded by the nucleic  
 CC acid molecule designated Gene 48 from the human cDNA clone HFCAT174  
 CC (deposited as clone ATCC 97899 and ATCC 209045). The gene can be used to  
 CC generate fusion proteins by linking to the gene to a human immunoglobulin  
 CC portion (e.g. AAVS9552) for increasing the stability of the fused  
 CC protein as compared to the human protein only. The invention relates to  
 CC 186 novel genes and their fragments (nucleic acid sequences: AAVS511-  
 CC V5812; amino acid sequences AAV4731-W5026) which are useful for  
 CC preventing, treating or ameliorating medical conditions e.g. by protein

CC or gene therapy. Also, pathological conditions can be diagnosed by  
 CC determining the amount of the new polypeptides in a sample or by  
 CC determining the presence of mutations in the new polynucleotides.  
 CC Specific uses are described for each of the 186 polynucleotides, based on  
 CC which tissues they are most highly expressed in (see AAVS9511 for  
 CC described uses)  
 XX  
 SQ Sequence 48 AA;  
 Query Match: 38.8%; Score 245; DB 2; Length 48;  
 Best Local Similarity: 100.0%; Pred. No. 1,7e-19;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 74 MGFPQWHLGNHVAPEVTSILLLFLMMLGVKGLLVGLVYVSHLSQR 121  
 DB 1 MGFPQWHLGNHVAPEVTSILLLFLMMLGVKGLLVGLVYVSHLSQR 48  
 RESULT 8  
 ABB68995  
 ID ABB68995 standard; protein; 48 AA.  
 AC ABB68995;  
 XX 20-JAN-2003 (first entry)  
 DT  
 XX Human polypeptide SEQ ID NO 1042.  
 DE  
 XX Human; genome mapping; gene therapy; food supplement; virus; fungus;  
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KW arthritis; cytostatic; immunomodulator; nociceptive; neuroprotective;  
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
 KW antiarthritic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200270539-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 05-MAR-2002; 2002WO-US005095.  
 XX  
 PR 05-MAR-2001; 2001US-00799451.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F,  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;  
 XX WPI, 2002-759812/82.  
 DR N-PSDB; AB211212.  
 DR  
 XX  
 PT New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative,  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
 PT or coagulation disorders.  
 XX  
 PS Claim 9; SEQ ID NO 1042; 1012pp + Sequence listing; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence selected from any of 948 sequences (AB21119-  
 CC AB21066) or their mature protein coding portion, active domain coding  
 CC protein or complementary sequences. The polynucleotides are useful for  
 CC identifying expressed genes or for physical mapping of human genome. The  
 CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
 CC markers, as a food supplement, for generating antibodies, in medical  
 CC imaging, screening and diagnostic assays and for treating cell-  
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,





XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC;  
PI Bednarik DR, Endress GA, Yu G, Ni U, Feng P, Young PE, Greene JM;  
PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Eshner R, Brewer LA;  
PI Moore PA, Shi Y, Latleur DM, Li Y, Zeng Z, Kyaw H;  
XX MPI; 2002-634796/68.  
DR N-PSDB; ABS73545.  
XX  
PT New isolated human secreted protein for diagnosing, preventing, treating  
PT or ameliorating medical conditions and used as a food additive or  
PT preservative.  
XX  
PS Example 1; SEQ ID NO 367; 129pp; English.  
XX  
XX The invention relates to an isolated protein that is one of 186 human  
CC secreted proteins, given in the specification, encoded by one of 309 cDNA  
CC sequences also given in the specification. The protein is used in a  
CC pharmaceutical composition used to prevent, treat or ameliorate a medical  
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. Disorders which are diagnosed or treated include  
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. The polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,  
CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
CC components. The present sequence represents one of the novel human  
CC secreted proteins of the invention. Note: This sequence did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from USPTO at seqdata.uspto.gov/sequence.htmlDocID=6420526B1  
XX  
SQ Sequence 48 AA;  
Query Match 38.8%; Score 245; DB 5; Length 48;  
Best Local Similarity 100.0%; Pred. No. 1.7e-19;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 74 MGFPQWHLGNHNPVTSTLLPLMLLGRGLLVGLVYVHLSQR 121  
Db 1 MGFPQWHLGNHNPVTSTLLPLMLLGRGLLVGLVYVHLSQR 48  
RESULT 10  
ID ABO34421 standard; protein; 48 AA.  
XX  
AC ABO34421;  
XX  
DT 22-SEP-2003 (first entry)  
XX  
DE Region of human secreted protein encoded by cDNA sequence #48.  
XX  
KM Human; secreted protein; hyperproliferative disorder; leukaemia;  
KM breast cancer; wound; reproductive disorder; blood-related disorder;  
KM hemophilia; thrombocytopenia; immunodeficiency; thymic hypoplasia;  
KM Wiskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;  
KM graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;  
KM viral infection; bacterial infection; fungal infection; AIDS; sepsis;  
KM renal infection; kidney failure; cardiovascular disorder; cytostatic;  
KM angina pectoris; cerebral ischaemia; congenital heart defect;  
KM respiratory disorder; neurological disorder; Alzheimer's disease;  
KM Parkinson's disease; inflammation; Crohn's disease; vulnerability;  
KM immunosuppressive; antibacterial; haemostatic; thrombolytic;  
KM anticoagulant; neuroprotective; thyromimetic; antiallergic;

KM antiasthmatic; virucide; fungicide; anti-HIV; nephrotoxic; antianginal;  
KM cerebroprotective; cardiant; nootropic; antiparkinsonian;  
KM antiinflammatory.  
XX  
OS Homo sapiens.  
XX  
XX US2003049618-A1.  
XX  
PD 13-MAR-2003.  
XX  
PF 16-MAR-2001; 2001US-00809391.  
XX  
XX 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043669P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 11-APR-1997; 97US-0047492P.  
PR 23-MAY-1997; 97US-0047500P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.  
PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047581P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.  
PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048684P.  
PR 06-JUN-1997; 97US-0048974P.  
PR 13-JUN-1997; 97US-0049610P.  
PR 08-JUL-1997; 97US-0051926P.  
PR 16-JUL-1997; 97US-0052874P.  
PR 18-AUG-1997; 97US-0055724P.



XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 20; SEQ ID NO 30992; 103bp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (II) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pcr\_sequences  
CC  
XX Sequence 114 AA;  
SQ  
Query Match 29.1%; Score 184; DB 4; Length 114;  
Best Local Similarity 44.4%; Pred. No. 2.6e-12;  
Matches 56; Conservative 4; Mismatches 38; Indels 28; Gaps 6;  
OY 1 MVRLLANGELVODDDPRVTRTTPGSGIRPSFRNGHAPGPGPPROQAGAR-LGA 59  
DB 1 MVRLLANGELVODRRP-----PSEDRP-----ATKRGSAVAPAKORCQGMVL 44  
OY 60 AQSEFNDLRLQTL-VNMGEPQWHLGNHAYEPY-----TSILLFLMMLGVRLILVGLV 112  
DB 45 LSSFFNDLKPAAGRTWAFRSHLR-----QPLVAGSDPSCILFLMMLGVRLILVGLV 100  
OY 113 YLVSHL 118  
DB 101 YLVSHL 106  
RESULT 12  
ABBS4669  
ID ABBS4669 standard; protein; 132 AA.  
XX  
XX ABB84669;  
AC  
XX 11-FEB-2003 (first entry)  
DT  
XX  
DE Human SECP-21 protein from clone 1265837CD1 SEQ ID 21.  
XX  
XX Secreted protein; SECP; human; antiarteriosclerotic; antiatherosclerotic;  
KM hepatocytic; cytosolic; anti-HIV; antiatherogenic; antiatherosclerotic; cancer;  
KM antidiabetic; antiinflammatory; neuroprotective; antiulcer;  
KM antineumatic; antidiabetic; cardiatic; hypotensive; gonadal dysgenesis;  
KM vasotropic; anticonvulsant; nootropic; immunosuppressive; pericarditis;  
KM antiparkinsonian; ophthalmological; cell proliferative disorder;  
KM arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; angina pectoris;  
KM autoimmune disorder; AIDS; Crohn's disease; multiple sclerosis; epilepsy;  
KM ulcerative colitis; cardiovascular disorder; myocardial infarction;  
KM Raynaud's disease; myocarditis; neurological disorder; cataract;  
KM Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
KM developmental disorder; Duchenne muscular dystrophy; antiporiatic;  
KM Becker muscular dystrophy; Cushing's syndrome.  
XX

OS Homo sapiens.  
XX  
XX WO200279441-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 29-MAR-2002; 2002MO-US009820.  
XX  
XX 30-MAR-2001; 2001US-0280527P.  
XX 06-APR-2001; 2001US-0282112P.  
XX 09-APR-2001; 2001US-0282702P.  
XX 13-APR-2001; 2001US-0283855P.  
XX 19-OCT-2001; 2001US-0343718P.  
XX 07-DEC-2001; 2001US-0339236P.  
XX 13-FEB-2002; 2002US-0357002P.  
XX  
XX (INCYTE) INCYTE GENOMICS INC.  
XX  
XX Baughn MR, Burford N, Ding L, Duggan BW, Elliott VS, Forsythe IU;  
XX Gandhi AR, Gietzen KU, Griffin JA, He A, Honchall CD, Ison CH;  
XX Lal PG, Lee EA, Lee S, Lu DM, Mason PM, Santhanala MM;  
XX Swarnakar A, Ramkumar J, Tang YT, Thangavelu K, Tran UK, Walla NK;  
XX Warren BA, Yao MG, Xu Y, Yue H;  
XX WPI; 2003-058429/05.  
XX N-PSDB; ABS57565.  
XX  
XX Novel human secreted protein useful for treating, preventing or  
PT diagnosing cancer, hepatitis, psoriasis, asthma, diabetes mellitus,  
PT anemia, epilepsy, cataract, Alzheimer's disease.  
XX  
XX Claim 76; Page 167; 188pp; English.  
XX  
XX This invention describes novel secreted proteins (SECP) which have  
CC antiarteriosclerotic, antiatherosclerotic, hepatotropic, cytosolic, anti  
CC -HIV, antiallergic, antiasthmatic, antidiabetic, antidiabetic, anti  
CC antiinflammatory, neuroprotective, antiulcer, antiparkinsonian, vasotropic,  
CC antineumatic, antidiabetic, cardiatic, hypotensive, anticonvulsant,  
CC nootropic, immunosuppressive, antiparkinsonian and ophthalmological  
CC activity. The polynucleotides and polypeptides of the invention can be  
CC used for diagnosing, treating or preventing cell proliferative disorder  
CC e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, cancer,  
CC autoimmune/inflammatory disorders e.g. acquired immunodeficiency syndrome  
CC (AIDS), allergies, asthma, anaemia, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, ulcerative colitis, psoriasis, rheumatoid arthritis,  
CC etc; cardiovascular disorder e.g. myocardial infarction, angina pectoris,  
CC hypertension, Raynaud's disease, myocarditis, pericarditis, etc;  
CC neurological disorders e.g. epilepsy, Huntington's disease, Parkinson's  
CC disease, Alzheimer's disease, Creutzfeldt-Jakob disease, etc; and  
CC developmental disorders e.g. Duchenne and Becker muscular dystrophy,  
CC cataract, gonadal dysgenesis, Cushing's syndrome, etc. The products of  
CC the invention can also be used for drug screening, proteome analysis,  
CC microarray creating knock-in humanised animals or transgenic animals to  
CC model human diseases, in somatic or germline gene therapy, to generate a  
CC transcript image of a tissue or cell type, for detecting differences in  
CC the chromosomal location due to translocation, inversion, etc., among  
CC normal, carrier or affected individuals, and as hybridization probes for  
CC mapping naturally occurring genomic sequences. ABB84649-ABB84673  
CC represent secreted proteins encoded by the CDNA's shown in ABS57545-  
CC ABS57569, described in the disclosure of the invention  
XX  
XX Sequence 132 AA;  
SQ  
Query Match 17.5%; Score 110.5; DB 6; Length 132;  
Best Local Similarity 31.9%; Pred. No. 0.00042;  
Matches 30; Conservative 17; Mismatches 34; Indels 13; Gaps 3;  
OY 26 RSIPROS-----FNRGAGAPGPGPPROQAGARLGAQSEFNDLRLQTLVNMGPQWHL 81  
DB 42 RQPRKESQVDDSDQNTGEPVGDYK-----KMGTL---FGLINKLIMGFTMTWF 92  
OY 82 GNHAYEPYTSILLFLMMLGVRLILVGLVYL 115



AC AAB42404;  
 XX 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF2168 polypeptide sequence SEQ ID NO:4336.  
 XX  
 KM Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KM vulnary; antipsoptic; antiparkinsonian; nootropic; neuroprotective;  
 KM anticonvulsant; osteopathic; antirheumatic; immunosuppressant; cardiac;  
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KM hypotensive; dermatological; immunosuppressive; antihypertensive;  
 KM antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;  
 KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease; AIDS;  
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KM cholesterol ester storage; systemic lupus erythematosus; infection;  
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KM bone damage; cartilage damage; antinflammatory disease; coagulation;  
 KM thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000MO-US008621.  
 XX  
 PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 PA  
 PI Shimkets RA, Leach M;  
 XX  
 XX WPI: 2000-602362/57.  
 DR N-PSDB; AAC76613.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 11; Page 3531-3532; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 CC antipsoptic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antirheumatic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antihypertensive; antianaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 XX Sequence 325 AA;

Query Match 12.7%; Score 80; DB 3; Length 325;  
 Best Local Similarity 35.5%; Pred. No. 2.8;  
 Matches 27; Conservative 13; Mismatches 30; Indels 6; Gaps 3;

QY 48 PROOAGARLGAAGSPFNDLNRQLVNM--GFPQWHLGNHAVEPVTSTLLFLMLMGVGR 105  
 DB 28 PFQALGALVNVQIFL--LLGQLVEVAVAKYTRDHVGSFMT--SQNLSTHLLILYVG 83  
 QY 106 ILLVGLVYLVSHLSOR 121  
 DB 84 LITFGYLVLSHVGGR 99

Search completed: August 24, 2004, 09:18:12  
 Job time : 61 secs

**This Page Blank (uspto)**

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:05 ; Search time 15 seconds  
(without alignments)  
775.945 Million cell updates/sec

Title: US-10-001-885-125

Sequence: 1 MVRILANGRIYQDDPRVRT.....GVKGLLVGLVYVLSLSQR 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 83.5  | 13.2        | 219    | 2 T46472 | hypothetical prote  |
| 2          | 81    | 12.8        | 475    | 2 A87396 | sensor histidine k  |
| 3          | 77.5  | 12.3        | 692    | 2 T21095 | hypothetical prote  |
| 4          | 76.5  | 12.1        | 1201   | 2 A83007 | hypothetical prote  |
| 5          | 76    | 12.0        | 159    | 2 UQ0145 | hypothetical 17.3K  |
| 6          | 74.5  | 11.8        | 1151   | 2 T18297 | zinc-finger protei  |
| 7          | 73    | 11.6        | 205    | 1 SAVIMC | delta large antige  |
| 8          | 73    | 11.6        | 651    | 2 T46050 | hypothetical prote  |
| 9          | 73    | 11.6        | 841    | 2 T01046 | hypothetical prote  |
| 10         | 72.5  | 11.5        | 195    | 1 SAVILI | delta large antige  |
| 11         | 72.5  | 11.5        | 207    | 2 UQ2167 | delta large antige  |
| 12         | 72.5  | 11.5        | 433    | 2 T36122 | probable integral   |
| 13         | 72    | 11.4        | 214    | 1 SAVIDV | delta large antige  |
| 14         | 71.5  | 11.3        | 552    | 1 QREBPT | aspartate chemorec  |
| 15         | 71.5  | 11.3        | 553    | 2 AC0746 | methyl-accepting c  |
| 16         | 71    | 11.2        | 345    | 2 A40616 | 34k antigen - Myco  |
| 17         | 71    | 11.2        | 345    | 2 A53138 | gas1 homolog - hum  |
| 18         | 71    | 11.2        | 618    | 2 T42664 | hypothetical prote  |
| 19         | 70.5  | 11.2        | 197    | 2 S22381 | hypothetical prote  |
| 20         | 70    | 11.1        | 146    | 2 E70557 | hypothetical prote  |
| 21         | 70    | 11.1        | 398    | 2 T34947 | hypothetical prote  |
| 22         | 70    | 11.1        | 2559   | 2 T09144 | probable guanine n  |
| 23         | 69.5  | 11.0        | 231    | 2 S75387 | hypothetical prote  |
| 24         | 69.5  | 11.0        | 269    | 2 T42919 | protein t10 - atel  |
| 25         | 69.5  | 11.0        | 447    | 2 T50705 | gamma-aminobutyrat  |
| 26         | 69.5  | 11.0        | 530    | 2 F87590 | serine protease     |
| 27         | 69    | 10.9        | 214    | 1 SAVIDN | delta large antige  |
| 28         | 69    | 10.9        | 255    | 2 B87585 | transcription regul |
| 29         | 69    | 10.9        | 283    | 2 G01926 | insulin promoter I  |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 69   | 10.9 | 484  | 2 AG2097 | serine/threonine k |
| 31 | 69   | 10.9 | 493  | 2 F75113 | NADH dehydrogenase |
| 32 | 69   | 10.9 | 495  | 2 G71019 | hypothetical prote |
| 33 | 69   | 10.9 | 1454 | 2 B84535 | probable retroelem |
| 34 | 68.5 | 10.8 | 148  | 2 S74589 | hypothetical prote |
| 35 | 68.5 | 10.8 | 195  | 2 SAVIMC | delta large antige |
| 36 | 68.5 | 10.8 | 195  | 2 S53112 | delta antigen - he |
| 37 | 68.5 | 10.8 | 206  | 2 T37094 | hypothetical prote |
| 38 | 68.5 | 10.8 | 454  | 1 C45340 | nucleosidase prote |
| 39 | 68.5 | 10.8 | 653  | 2 T34356 | hypothetical prote |
| 40 | 68   | 10.8 | 530  | 2 T48627 | hypothetical prote |
| 41 | 67.5 | 10.7 | 525  | 2 T44445 | chitinase (EC 3.2. |
| 42 | 67.5 | 10.7 | 1747 | 2 A54121 | collagen alpha-4 c |
| 43 | 67   | 10.6 | 260  | 2 S22373 | proline-rich prote |
| 44 | 67   | 10.6 | 914  | 2 T55220 | hypothetical prote |
| 45 | 67   | 10.6 | 1121 | 2 A82809 | exodeoxyribonuclea |

# ALIGNMENTS

RESULT 1  
T46472  
hypothetical protein DKFP434N2030.1 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C/Accession: T46472  
R/Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, U.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A/Reference number: Z23034  
A/Accession: T46472  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-219 <AAA>  
A/Cross-references: EMBL:AL137488  
A/Experimental source: adult testis; clone DKFP434N2030  
C/Genetics:  
A/Note: DKFP434N2030.1

Query Match 13.2%; Score 83.5; DB 2; Length 219;  
Best Local Similarity 36.5%; Pred. No. 0.56;  
Matches 23; Conservative 11; Mismatches 24; Indels 5; Gaps 2;

QY 41 APPGPGPRCOAGARLGAOSPFNDLNRQLVNMGFPMHLGNHAYEPTSLILFLMM 100  
DB 9 APPRAGSVRRRG-RLNALRLPC---RPGSVGFPEGVWGSSVRELVVLFVVR 63  
QY 101 LGV 103  
DB 64 EGI 66

RESULT 2  
A87396  
sensor histidine kinase [imported] - Caulobacter crescentus  
C/Species: Caulobacter crescentus  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C/Accession: A87396  
R/Heiman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Taub, M.T.; DeBoy, R.T.; Dodson, R.U.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, U.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of Caulobacter crescentus.  
A/Reference number: A87249; MUID:21173698; PMID:11259647  
A/Accession: A87396  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-475 <STO>  
A/Cross-references: GB:AE005673; NID:g13422503; PIDN:AAK23165.1; GSPDB:GN00148  
C/Genetics:  
A/Gene: CC1181

Query Match 12.8%; Score 81; DB 2; Length 475;



Best Local Similarity 29.0%; Pred. No. 2.2;  
Matches 36; Conservative 9; Mismatches 39; Indels 40; Gaps 7;

QY 9 EIVODDDPRV---RTTQPPRGSI-----RQSFNRGHGAPGGPGRQQA---53  
Db 104 ELQDNGRFRFVRRTPTTPRPDAPPGAGNCRPAKPEFRREG-GAGCGDNARSQARD 162  
QY 54 -----GARLGAAGSPENDLNRLV-----NMGFPQHLGNHAEVPTSLILFLMLTG 102  
Db 163 EPLIFGDFKVGVRQSD-----GRWLVEPKPTLRFDSSWQ-----QRILLILLISVIA 209  
QY 103 VRGL 106  
Db 210 VTPL 213

## RESULT 3

hypothetical protein F18H3.3a - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C/Accession: T21095

R/Coles, L.  
submitted to the EMBL Data Library, July 1995

A/Reference number: Z19373

A/Accession: T21095

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-692 <WTL>

A/Cross-references: EMBL:Z50110; PTDN:CA90444.1; GSPDB:GN0028; CESP:F18H3.3a

A/Experimental source: clone F18H3

C/Genetics:

A/Map position: X

A/Intons: 111/1; 215/2; 469/3; 552/3

C/Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Query Match 12.3%; Score 77.5; DB 2; Length 692;  
Best Local Similarity 13.9%; Pred. No. 7;  
Matches 37; Conservative 9; Mismatches 43; Indels 27; Gaps 7;

QY 12 QDDPRVTRTT-OPRGSIRQSFNRGHGAP-PGPG-----PQQQAGARLGAAGSP 63  
Db 511 QQGRPPMRITVDGRPOYGMARPPA-----GAPRGVGVGMGAPTRQGPAPRGAPOKR 564  
QY 64 P-DNLNQLVNMGFPQHLGNHAEVPTSLILFLMLTGVRGLLVGLVYVSHL 118  
Db 565 FYGQPPRQOQ---PQGHSGQPPAQPOQ-----QGGQSGIVIHGQETLISHM 608

## RESULT 4

hypothetical protein PA514 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C/Accession: A83007

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
J.; Lory, S.; Olson, M.V.

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
gen

A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: A83007

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1201 <STO>

A/Cross-references: GB:AE004924; GB:AE004091; NID:99951407; PTDN:AA08499.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Map position: X

A/Intons: 111/1; 215/2; 469/3; 552/3

C/Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Matches 29; Conservative 10; Mismatches 45; Indels 25; Gaps 4;

QY 16 PAVRTTQPPRGSIROSFFNRGHGAPGGPGRQQAARLGAAGSPFNDLNRLVNM 75  
Db 118 PAVEAAPPPRPPEPVQIERQAPQAPAPR-----AAPAPPREPSLVERG 168  
QY 76 PPO--WHLGNHAEVPTSLILFLMLTGVRGLLVGLVYVSHLSOR 121  
Db 169 FAAGKWLFGGNTV-----LRIGV-VLLFTGLAFILRYASER 204

## RESULT 5

hypothetical 17.3K protein - Pseudomonas aeruginosa

C/Species: Pseudomonas aeruginosa

C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 07-Jun-1996

C/Accession: J00145

R/Kato, J.; Chu, L.; Kitano, K.; Devault, J.D.; Kimbara, K.; Chakrabarty, A.W.; Misra, T.

Gene 84, 31-38, 1989

A/Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in Pse

A/Reference number: U00132; MUID:90108714; PMID:2514124

A/Accession: J00145

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-159 <KAT>

A/Note: 50-Met could also be the initiator

C/Genetics:

A/Start codon: GTG

Query Match 12.0%; Score 76; DB 2; Length 159;  
Best Local Similarity 32.9%; Pred. No. 2.3;  
Matches 23; Conservative 3; Mismatches 18; Indels 26; Gaps 4;

QY 26 RGSIRQSFNRGHGAPG-----GP-----GPRQQAARLGAAGSPFNDLNRLV 72  
Db 33 FAGVPARS-RRRGHGRPGWADRSAGETAGLPAQPRHLHAGSLPSSDP-----82  
QY 73 NMGFPQHLG 82  
Db 83 ---GPAMHKG 89

## RESULT 6

zinc-finger protein FOG-2 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C/Accession: T18297

R/Revesian, S.G.; Deconinck, A.E.; Cantor, A.B.; Rieff, H.I.; Fujiwara, Y.; Corfas, G.; C

Proc. Natl. Acad. Sci. U.S.A. 96, 950-955, 1999

A/Title: FOG-2: A novel GATA-family cofactor related to multitype zinc-finger proteins F

A/Reference number: Z18866; MUID:99128343; PMID:9927674

A/Accession: T18297

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1151 <TEV>

A/Cross-references: EMBL:AF107306; NID:94206772; PTD:94206773; PTDN:AA012182.1

C/Genetics:

A/Map position: X

A/Intons: 111/1; 215/2; 469/3; 552/3

C/Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Query Match 11.8%; Score 74.5; DB 2; Length 1151;  
Best Local Similarity 23.3%; Pred. No. 23;  
Matches 27; Conservative 19; Mismatches 33; Indels 37; Gaps 5;

QY 4 ILANGELVODDDPRVTRTTQPPRGSIROSFFNRGHGAP-----PGGPGP 48  
Db 1042 VIVVGGKQDPRP-----TANQGENISQNTQHDGHSFWSISENPLANENVSGIFCA 1097  
QY 49 RQQAARL---GAAGP-----FNDLNRLVNMGFPQHLGNHAEV 89  
Db 1098 EEOLSIAKGVNGASQAPSSGKXCRLCDIOFNINLNFITKRF---YCSSHAABEV 1150

RESULT 7  
 SAVLMC  
 delta large antigen - hepatitis delta virus (woodchuck isolate)  
 N:Alternate names: HDag  
 C:Species: hepatitis delta virus  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 01-Aug-1997  
 C:Accession: A30054  
 R:Kuo, M.Y.P.; Goldberg, J.; Coates, L.; Mason, W.; Gerin, J.; Taylor, J.  
 J. Virol. 62: 1955-1961, 1988  
 A:Title: Molecular cloning of hepatitis delta virus RNA from an infected woodchuck liver  
 A:Reference number: A30054; MUID:88215010; PMID:3367426  
 A:Accession: A30054  
 A:Molecule type: genomic RNA  
 A:Residues: 1-205 <K10>  
 A:Cross-references: GB:M21012  
 C:Comment: This virus is a replication-defective hepatitis B virus.  
 C:Superfamily: hepatitis delta virus large antigen  
 C:Keywords: core protein  
 F/2-205/Product: delta large antigen #status predicted <MAT>

Query Match 11.6%; Score 73; DB 1; Length 205;  
 Best Local Similarity 30.7%; Pred. No. 5.8;  
 Matches 23; Conservative 8; Mismatches 32; Indels 13; Gaps 4;

QY 10 IVODDDPRVATTTPPGSGIPRQSFNRGHPGAPGPGPRQOAGARLGAASPFNDLNR 69  
 DB 133 LTEDERERKRRVAGPVGCV--NPLEGSGRGAAPGGGFVNLQ-----GVPESPFRSGE 184

QY 70 QI--VVMGFPQNL 81  
 DB 185 GLDIRNGGFP-WDI 198

RESULT 8  
 T46050  
 hypothetical protein T16K5.190 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C:Accession: T46050  
 R:Reeger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke, submitted to the Protein Sequence Database, January 2000  
 A:Reference number: Z23015  
 A:Accession: T46050  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-651 <R1E>  
 A:Cross-references: EMBL:A1132965  
 A:Experimental source: cultivar Columbia; BAC clone T16K5  
 C:Genetics:  
 A:Map position: 3  
 A:Insertions: 52/1; 83/1; 153/2; 209/1; 257/3; 278/3; 309/3; 363/3; 424/1; 479/1; 501/1; 54  
 A:Note: T16K5.190

Query Match 11.6%; Score 73; DB 2; Length 651;  
 Best Local Similarity 27.0%; Pred. No. 18;  
 Matches 17; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

QY 15 DPRVATTTPPGSGIPRQSFNRGHPGAPGPGPRQOAGARLGAASPFNDLNRQLVNM 74  
 DB 532 DPGHPVAPPPGCGPRKEGTPRAGYPPAPYPPQYQAVPPAGYPPPGQGGGQVPA 591

QY 75 GFP 77  
 DB 592 GYP 594

RESULT 9  
 T01046  
 hypothetical protein YUP8H12R.29 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 22-Oct-1999  
 C:Accession: T01046  
 R:Theologis, A.; Vysotskaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwan

Oefner, P.; Davis, R.M.  
 submitted to the EMBL Data Library, May 1998  
 A:Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.  
 A:Reference number: Z14227  
 A:Accession: T01046  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-841 <T1E>  
 A:Cross-references: EMBL:AC002986; NID:g2494106; PID:g3152568; GSPDB:GN00059; ATSP:YUP8H  
 C:Genetics:  
 A:Gene: ATSP:YUP8H12R.29  
 A:Map position: 1  
 A:Insertions: 74/3; 93/3; 164/1

Query Match 11.6%; Score 73; DB 2; Length 841;  
 Best Local Similarity 35.8%; Pred. No. 24;  
 Matches 24; Conservative 8; Mismatches 19; Indels 16; Gaps 5;

QY 29 IPROSFFNRGHPGAPGPGPRQ-----QOAGARLGAAG-SPFNDLNRQLVNM--GF 76  
 DB 241 VPKSFSYSY---PPPGSIPQRLGHPNIPYQSGPGQGSFNPFPNLQPLPSMHGS 297

QY 77 PGMHGN 83  
 DB 298 PQ-HTGN 303

RESULT 10  
 SAVLMC  
 delta large antigen - hepatitis delta virus (strain Lebanon-1)  
 N:Alternate names: HDag  
 C:Species: hepatitis delta virus  
 A:Note: host Homo sapiens (man)  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 01-Aug-1997  
 C:Accession: A40247  
 R:Lee, C.M.; Bih, F.Y.; Chao, Y.C.; Govindarajan, S.; Lai, M.M.C.  
 Virology 188, 265-273, 1992  
 A:Title: Evolution of hepatitis delta virus RNA during chronic infection.  
 A:Reference number: A40247; MUID:92230225; PMID:1566577  
 A:Accession: A40247  
 A:Molecule type: genomic RNA  
 A:Residues: 1-195 <L1E>  
 A:Cross-references: GB:M64917  
 C:Comment: This virus is a replication-defective hepatitis B virus.  
 C:Superfamily: hepatitis delta virus large antigen  
 C:Keywords: core protein  
 F/2-195/Product: delta large antigen #status predicted <MAT>

Query Match 11.5%; Score 72.5; DB 1; Length 195;  
 Best Local Similarity 31.5%; Pred. No. 6.1;  
 Matches 23; Conservative 8; Mismatches 31; Indels 11; Gaps 3;

QY 8 GRIVODDPRVATTTPPGSGIPRQSFNRGHPGAPGPGPRQOAGARLGAASPFNDL 67  
 DB 131 GRLTEDEKRRKRRVAGPPTGVNPLEGGR--GAPGGGFVPSMQ-----GVPESPFRH 182

QY 68 NRQL--VVMGFP 77  
 DB 183 GEGLDARDGRFP 195

RESULT 11  
 JQ2167  
 delta large antigen - hepatitis delta virus (strain CAR)  
 N:Alternate names: HDag  
 C:Species: hepatitis delta virus  
 C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 07-May-1999  
 C:Accession: JQ2167  
 R:Tang, J.R.; Hantz, O.; Vitvitski, L.; Lamelin, J.P.; Parana, R.; Cova, L.; Lesbordes, U.; Gen. Virol. 74, 1827-1835, 1993  
 A:Title: Discovery of a novel point mutation changing the HDag expression of a hepatitis A:Reference number: JQ2167; MUID:93389437; PMID:8376962  
 A:Accession: JQ2167

A:Molecule type: mRNA  
A:Residues: 1-207 <TAN>  
A:Experimental source: isolate CAR  
A>Note: the authors translated the codon CGA for residue 85 as nothing and AGA for residue 86 as nothing.  
C:Comment: This protein is one of hepatitis delta virus specific proteins.  
C:Superfamily: hepatitis delta virus large antigen  
C:Keywords: core protein, lipoprotein, methylated carboxyl end, prenylated cysteine  
F/204/Binding site: farnesyl (Cys) (covalent) #status predicted  
F/204/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 11.5%; Score 72.5; DB 2; Length 207;  
Best Local Similarity 30.6%; Pred. No. 6.5; Indels 11; Gaps 3;  
Matches 22; Conservative 8; Mismatches 31;

QY 10 IVQDDPRVTRTTTPRPGSIPRQSFNRGAGAPGPGPQQAGARLGAAGSPFNDLN 69  
DB 126 LTEDERRERRAAPVGV--NPLEGSGRAGPGGFPVPLQ-----GVPSFRTGB 177  
QY 70 QL---VNMGPFG 78  
DB 178 GLDIRGNQGFPG 189

## RESULT 12

T36122  
Probable integral membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T36122

R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, April 1999  
A:Reference number: Z21597  
A:Accession: T36122  
A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA  
A:Residues: 1-433 <MUR>  
A:Cross-references: EMBL:AL049707; PIDN:CA841288.1; GSPDB:GN00070; SCODEB:SCB15.19  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCODEB:SCB15.19

Query Match 11.5%; Score 72.5; DB 2; Length 433;  
Best Local Similarity 31.3%; Pred. No. 14;

Matches 26; Conservative 9; Mismatches 31; Indels 17; Gaps 5;

QY 11 VQDDPRVTRTTTPRPGSIPRQSFNRGAGAPGPGPQQAGARLGAAGSPFNDLN 68  
DB 241 LQETDPQAPAGEQ--RAELVKGPEADRPRAEPSPGAGSEDPGSPWPN--T 296

QY 69 RQLVNMG-----FPQWH 80  
DB 297 RQVAYQGYDPSPYDPGQVQWY 319

## RESULT 13

SAV1DV  
delta large antigen - hepatitis delta virus (chimpanzee isolate)

N:Alternate names: HDAG  
C:Species: hepatitis delta virus

C>Date: 30-Sep-1987 #sequence\_revision 31-Dec-1990 #text\_change 01-Aug-1997  
C:Accession: A26176; A30028

R:Wang, K.S.; Choo, Q.L.; Weiner, A.J.; Ou, J.H.; Najarian, R.C.; Thayer, R.M.; Mullens  
Nature 323, 508-514, 1986  
A:Title: Structure, sequence and expression of the hepatitis delta (delta) viral genome.  
A:Reference number: A26176; PMID:87014821; PMID:3762705

A:Accession: A26176  
A:Molecule type: genomic RNA  
A:Residues: 1-161 <MAN>  
A:Cross-references: GB:X04451

A>Note: the nucleotide sequence contains a frameshift error in codon 162 that has been corrected.  
R:Wang, K.S.; Choo, Q.L.; Weiner, A.J.; Ou, J.H.; Najarian, R.C.; Thayer, R.M.; Mullens  
Nature 323, 456, 1987  
Reference number: X00078

A:Contents: corrections  
A:Accession: A30028  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: genomic RNA  
A:Residues: 1-214 <MA>  
A:Cross-references: GB:M21012

A>Note: this corrects a frameshift error in codon 162 of the published sequence  
C:Comment: This virus is a replication-defective hepatitis B virus.  
C:Superfamily: hepatitis delta virus large antigen  
C:Keywords: core protein, lipoprotein, methylated carboxyl end, prenylated cysteine  
F/2-211/Product: delta large antigen #status predicted <MA>  
F/211/Binding site: farnesyl (Cys) (covalent) #status predicted  
F/211/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 11.4%; Score 72; DB 1; Length 214;  
Best Local Similarity 32.9%; Pred. No. 7.6;  
Matches 25; Conservative 8; Mismatches 29; Indels 14; Gaps 5;

QY 10 IVQDDPRVTRTTTPRPGSIPRQSFNRGAGAPGPGPQQAGARLGAAGSPFNDLN 68  
DB 133 LTEDERRERRAAPVGVPLR---GSGRAGGPGFPVPLQ-----GVPSFRTGB 183

QY 69 RQL---VNMGPFG 81  
DB 184 GLDIRGNQGFPG 198

## RESULT 14

Q8BBDT

aspartate chemoreceptor protein - Salmonella typhimurium

N:Alternate names: methyl-accepting chemotaxis protein II (MCP-II)  
C:Species: Salmonella typhimurium  
C>Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 15-Aug-1997  
C:Accession: A03441; A29053

R:Russo, A.F.; Koshland Jr., D.E.  
Science 220, 1016-1020, 1983

A:Title: Separation of signal transduction and adaptation functions of the aspartate receptor.  
A:Reference number: A03441; PMID:83197387; PMID:6302843  
A:Accession: A03441

A:Molecule type: DNA  
A:Residues: 1-552 <RUS>

A>Note: residues 294-296, 301-302, 308-309, 483-484, 490-491, and 499-501 are potential  
R:Mowbray, S.L.; Foster, D.L.; Koshland Jr., D.E.  
J. Biol. Chem. 260, 11711-11718, 1985

A:Title: Chemical fragments identified with domains of the aspartate chemoreceptor.  
A:Reference number: A29053; PMID:8608217; PMID:2295347

A:Accession: A29053  
A:Molecule type: protein  
A:Residues: 1-245, 'E', 247-552 <MOW>

C:Comment: This protein responds to changes in aspartate concentration in the environment  
C:Genetics:  
A:Gene: tar

C:Superfamily: methyl-accepting chemotaxis protein  
C:Keywords: chemotaxis; methylated amino acid; sensory transduction; transmembrane protein  
F/295/309/Modified site: glutamate methyl ester (Glu) (by cheB-dependent deamidation and  
F/302/491/500/Modified site: glutamate methyl ester (Glu) #status predicted

Query Match 11.3%; Score 71.5; DB 1; Length 552;  
Best Local Similarity 28.2%; Pred. No. 22;  
Matches 22; Conservative 15; Mismatches 28; Indels 13; Gaps 3;

QY 50 QQAGARLGAAGSPFNDLNQVNMG-----FPQWHGNAVPEVPSILFLPMWL--G 102  
DB 158 QNALGKALGVYARSEULVQTPQSAHDYRFQWQGL-----VLAVALVILVYVWF 211

QY 103 VQGLLVGLVYVSHSQ 120  
DB 212 IRRALNPLARVTHIRE 229

## RESULT 15

100746



**This Page Blank (uspio)**

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:04 ; Search time 9 Seconds

(Without alignments)  
700.054 Million cell updates/sec

Title: US-10-001-885-125

Sequence: 1 MVRILNGEIVQDDPRVET.....GVRGLLVGLVYLVSHTSOR 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 75    | 11.9        | 735    | 1     | ABC8_HUMAN  |
| 2          | 75    | 11.6        | 735    | 1     | ABC8_HUMAN  |
| 3          | 72.5  | 11.5        | 195    | 1     | AAAT_HUMAN  |
| 4          | 72    | 11.4        | 195    | 1     | AAAT_HUMAN  |
| 5          | 72    | 11.4        | 769    | 1     | AD11_HUMAN  |
| 6          | 71.5  | 11.3        | 490    | 1     | FA10_RABIT  |
| 7          | 71.5  | 11.3        | 553    | 1     | MCP2_SALTY  |
| 8          | 71    | 11.2        | 298    | 1     | 34XD_MYCPA  |
| 9          | 71    | 11.2        | 345    | 1     | GAS1_HUMAN  |
| 10         | 71    | 11.2        | 478    | 1     | CTP8_HUMAN  |
| 11         | 70.5  | 11.1        | 481    | 1     | CTP8_HUMAN  |
| 12         | 70    | 11.1        | 1054   | 1     | S24A_MOUSE  |
| 13         | 69    | 10.9        | 214    | 1     | AAAT_HUMAN  |
| 14         | 69    | 10.9        | 283    | 1     | IPF1_HUMAN  |
| 15         | 69    | 10.9        | 283    | 1     | IPF1_HUMAN  |
| 16         | 68.5  | 10.8        | 195    | 1     | AAAT_HUMAN  |
| 17         | 68.5  | 10.8        | 195    | 1     | AAAT_HUMAN  |
| 18         | 68.5  | 10.8        | 195    | 1     | AAAT_HUMAN  |
| 19         | 68.5  | 10.8        | 1070   | 1     | PCB8_HUMAN  |
| 20         | 68    | 10.8        | 839    | 1     | APB1_RAT    |
| 21         | 67    | 10.6        | 2191   | 1     | POLG_HUMAN  |
| 22         | 66.5  | 10.5        | 255    | 1     | LYOX_HUMAN  |
| 23         | 66.5  | 10.5        | 420    | 1     | LYOX_HUMAN  |
| 24         | 66.5  | 10.5        | 424    | 1     | LYOX_HUMAN  |
| 25         | 66.5  | 10.5        | 738    | 1     | AB10_HUMAN  |
| 26         | 66.5  | 10.5        | 738    | 1     | AB10_HUMAN  |
| 27         | 66.5  | 10.5        | 1125   | 1     | IF2_HUMAN   |
| 28         | 66    | 10.4        | 283    | 1     | IPF1_RAT    |
| 29         | 66    | 10.4        | 628    | 1     | TEST_RAT    |
| 30         | 66    | 10.4        | 705    | 1     | HS70_HUMAN  |
| 31         | 65.5  | 10.4        | 1505   | 1     | CUT1_HUMAN  |
| 32         | 65.5  | 10.4        | 286    | 1     | PMF3_MOUSE  |
| 33         | 65.5  | 10.4        | 407    | 1     | AMBN_MOUSE  |

|    |      |      |      |   |            |
|----|------|------|------|---|------------|
| 34 | 65.5 | 10.4 | 1462 | 1 | NC02_MOUSE |
| 35 | 65   | 10.3 | 362  | 1 | HRCA_BRAJA |
| 36 | 65   | 10.3 | 385  | 1 | ITR5_BOVIN |
| 37 | 65   | 10.3 | 461  | 1 | SYT_SALTY  |
| 38 | 65   | 10.3 | 461  | 1 | SYT_SALTY  |
| 39 | 65   | 10.3 | 606  | 1 | ASAF_PIG   |
| 40 | 65   | 10.3 | 666  | 1 | ENV_MLVHO  |
| 41 | 65   | 10.3 | 674  | 1 | CA1A_BOVIN |
| 42 | 65   | 10.3 | 1596 | 1 | MAV_DROME  |
| 43 | 64.5 | 10.2 | 297  | 1 | PCW_BROWH  |
| 44 | 64.5 | 10.2 | 361  | 1 | SSB3_RAT   |
| 45 | 64.5 | 10.2 | 388  | 1 | SSB3_HUMAN |

#### ALIGNMENTS

RESULT 1  
ABC8\_HUMAN STANDARD; PRT; 735 AA.  
ID AC Q9NUT2; 095787;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE ATP-binding cassette, sub-family B, member 8, mitochondrial precursor  
DE (Mitochondrial ATP-binding cassette 1) (M-ABC1).  
GN ABC8 OR MAB1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT), AND MUTAGENESIS.  
RX MEDLINE=99096930; PubMed=9878413;  
RA Hogue D.L., Liu L., Ling V.;  
RA "Identification and characterization of a mammalian mitochondrial ATP-  
binding cassette membrane protein.";  
RL J. Mol. Biol. 285:379-389 (1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM LONG).  
RC TISSUE=Placenta;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Iehibashi T., Fujimori K.,  
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,  
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K.,  
RA Masuko Y., Kanehori K.;  
RL "NEO human cDNA sequencing project.";  
RN Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RP VARIANT ILE-152.  
RX MEDLINE=21686803; PubMed=11829140.  
RA Saito S., Iida A., Sekine A., Miura Y., Ogawa C., Kawauchi S.,  
RA Higuchi S., Nakamura Y.;  
RT "Three hundred twenty-six genetic variations in genes encoding nine  
members of ATP-binding cassette, subfamily B (ABC8/MDR/TFP), in the  
Japanese population.";  
RL J. Hum. Genet. 47:38-50(2002).  
CC -!- SUPNUT: Monomer (Potential).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
inner membrane.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long;  
CC IsoId=Q9NUT2-1; Sequence=Displayed;  
CC Name=Short;  
CC IsoId=Q9NUT2-2; Sequence=VSP\_000026;  
CC Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Ubiquitous.  
CC -!- SIMILARITY: Belongs to the ABC transporter family. MDR subfamily.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL; AF047690; AAD15748.1; -.
CC EMBL; AK002018; BAA92038.1; -.
CC HSSP; P13569; INBD.
CC Genew; HGNC:49; ABCB8.
CC MIM; 605464; -.
CC GO; GO:0005624; C:membrane fraction; TAS.
CC GO; GO:0005740; C:mitochondrial membrane; TAS.
CC GO; GO:0005524; F:ATP binding; TAS.
CC GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; TAS.
CC GO; GO:0006810; P:transport; TAS.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001140; ABC_TM_transp.
CC InterPro; IPR003439; ABC_transporter.
CC Pfam; PF00664; ABC_membrane; 1.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF00006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00929; ABC_TMLF; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
CC ATP-binding; Transmembrane; Transport; Mitochondrion; Inner membrane;
CC Transmembrane; Alternative splicing; Polymorphism.
CC TRANSIT ? 735
CC CHAIN ? 735
CC TRANSMEM 86 102
CC TRANSMEM 145 165
CC TRANSMEM 196 216
CC TRANSMEM 296 316
CC TRANSMEM 383 403
CC NP_BIND 507 514
CC VASPLIC 33 49
CC VARIANT 152 152
CC MUTAGEN 512 513
CC CONFLICT 174 174
CC CONFLICT 239 239
CC CONFLICT 278 278
CC CONFLICT 326 326
CC CONFLICT 347 347
CC CONFLICT 365 365
CC SEQUENCE 735 AA; 79947 MW; 879519858F7E0C7A6 CRC64;

Query Match 11.6%; Score 75; DB 1; Length 735;
Best Local Similarity 35.7%; Pred. No. 9.9;
Matches 25; Conservative 13; Mismatches 26; Indels 6; Gaps 3;

Db 157 GAALVNVQIPL--LLGQLVNVAKYTRDVGSPFMT--SQNLSTHLLILYGVGLLFTG 212
QY 112 VYIVSHLSQR 121
QY 213 LVLLSHVGER 222

RESULT 2
AANT_HDVMO STANDARD; PRT; 205 AA.
ID AANT_HDVMO
AC P29997;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
CC Delta antigen.
CC Hepatitis delta virus (isolate Woodchuck) (HDV).
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC NCBI_TaxID=31764;
CC [1]
CC MEDLINE=86215010; PubMed=3367426;
CC Kuo W.Y.F., Goldberg J., Coates U., Mason W., Gerin J., Taylor J.;
CC "Molecular cloning of hepatitis delta virus RNA from an infected
CC woodchuck liver: sequence, structure, and applications.";
CC RT J. Virol. 62:1855-1861(1988).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: Phosphorylated.
CC -1- DISEASE: BINDS SPECIFICALLY TO ANTISERA FROM PATIENTS WITH CHRONIC
CC HEPATITIS DELTA VIRAL INFECTIONS.
CC -1- MISCELLANEOUS: Delta antigen binds specifically to HDV RNA.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL; M21012; AAA45723.1; ALT_TERM.
CC PIR; A30054; SAVIMC.
CC InterPro; IPR002506; HDV_ag.
CC Pfam; PF01517; HDV_ag; 1.
CC ProDom; PD002887; HDV_ag; 1.
CC Antigen; RNA-binding; Nuclear protein; Phosphorylation.
CC SEQUENCE 205 AA; 23047 MW; 8B33C73B7872B98 CRC64;

Query Match 11.6%; Score 73; DB 1; Length 205;
Best Local Similarity 30.7%; Pred. No. 4.1;
Matches 23; Conservative 8; Mismatches 32; Indels 12; Gaps 4;

Db 10 IVODDDPRVETTTQPPRGSIIRGAGPGGPGPRQQAAGTGAOSPFNDLR 69
QY 133 LTBEDEBRERVAAPPGVG--NPLEGSRKAPGGGVPMNQ-----GVPESEFSRTGE 164
QY 70 QL--VNMGPQWHL 81
QY 185 GLDIRGNQGP-MDI 198

RESULT 3
AANT_HDVLI STANDARD; PRT; 195 AA.
ID AANT_HDVLI
AC P29833;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
CC Delta antigen.
CC Hepatitis delta virus (isolate Lebanon-1) (HDV).
CC Viruses; Deltavirus.
CC NCBI_TaxID=31763;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=92230225; PubMed=1566577;
CC Lee C.M., Bin F.Y., Chao Y.C., Govindarajan S., Lai M.M.C.;
CC "Evolution of hepatitis delta virus RNA during chronic infection.";
CC RT Virology 188:265-273(1992).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: Phosphorylated.
CC -1- DISEASE: BINDS SPECIFICALLY TO ANTISERA FROM PATIENTS WITH CHRONIC
CC HEPATITIS DELTA VIRAL INFECTIONS.
CC -1- MISCELLANEOUS: Delta antigen binds specifically to HDV RNA.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```



CC EMBL; M84917; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A40247; SAVILL.  
 DR InterPro; IPR002506; HDV\_ag.  
 DR Pfam; PF01517; HDV\_ag; 1.  
 DR ProDom; PD002887; HDV\_ag; 1.  
 DR Antigen; RNA-binding; Nuclear protein; Phosphorylation.  
 SQ SEQUENCE 195 AA; 21846 MW; 288929CS95B5D41B CRC64;  
 Query Match 11.5%; Score 72.5; DB 1; Length 195;  
 Best Local Similarity 31.5%; Pred. No. 4.3;  
 Matches 23; Conservative 8; Mismatches 31; Indels 11; Gaps 3;  
 QY 8 GEIVODDPRVYRTTTPRGSIPRQSFNKGHAPGPGPGPQQAGARLGAASPFNDL 67  
 DB 131 GRLTEDEKRRRVAGPFTGGVNLGGQR--GAPGGGFVPSNQ-----GVPESPFRRH 182  
 QY 68 NRQI---VNMGFP 77  
 DB 183 GEGLDARGDRGFP 195  
 RESULT 4  
 AANT\_HDVIT STANDARD; PRT; 214 AA.  
 AC P06934;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Delta antigen (Alpha antigen).  
 OS Hepatitis delta virus (isolate Italian) (HDV).  
 OC Viruses; Deltavirus.  
 OX NCBI\_TaxID=10423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=67014821; PubMed=3762705;  
 RA Wang K.S., Choo Q.L., Weiner A.J., Ou J.H., Najarian R.C.,  
 RA Thayer R.M., Mullendach G.T., Denniston K.J., Gerin J.L., Houghton M.;  
 RT "Structure, sequence and expression of the hepatitis delta (delta)  
 RT viral genome.";  
 RL Nature 323:508-514(1986).  
 RN [2]  
 RP REVISIONS TO 162-214.  
 RA Wang K.S., Choo Q.L., Weiner A.J., Ou J.H., Najarian R.C.,  
 RA Thayer R.M., Mullendach G.T., Denniston K.J., Gerin J.L., Houghton M.;  
 RL Nature 328:456-456(1987).  
 RN [3]  
 RP SUBCELLULAR LOCATION, AND PHOSPHORYLATION.  
 RX MEDLINE=88230601; PubMed=3373572;  
 RA Chang M.F., Baker S.C., Soe L.H., Kamahora T., Keck J.G., Makino S.,  
 RA Govindarajan S., Lai M.M.C.;  
 RT "Human hepatitis delta antigen is a nuclear phosphoprotein with RNA-  
 RT binding activity.";  
 RL J. Virol. 62:2403-2410(1988).  
 RN [4]  
 RP RNA-BINDING SPECIFICITY.  
 RX MEDLINE=90347794; PubMed=2200884;  
 RA Lin J.-H., Chang M.F., Baker S.C., Govindarajan S., Lai M.M.C.;  
 RT "Characterization of hepatitis delta antigen: specific binding to  
 RT hepatitis delta virus RNA.";  
 RL J. Virol. 64:4051-4058(1990).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- PPM: Phosphorylated.  
 CC -1- DISEASE: BINDS SPECIFICALLY TO ANTISERA FROM PATIENTS WITH CHRONIC  
 CC HEPATITIS DELTA VIRAL INFECTIONS.  
 CC -1- MISCELLANEOUS: Delta antigen binds specifically to HDV RNA.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC EMBL; X04451; CAA28050.1; ALT\_SEQ.  
 DR PIR; A26176; SAVILDV.  
 DR InterPro; IPR002506; HDV\_ag.  
 DR Pfam; PF01517; HDV\_ag; 1.  
 DR ProDom; PD002887; HDV\_ag; 1.  
 DR Antigen; RNA-binding; Nuclear protein; Phosphorylation.  
 SQ SEQUENCE 214 AA; 24078 MW; 83221B6EED0A2688 CRC64;  
 Query Match 11.4%; Score 72; DB 1; Length 214;  
 Best Local Similarity 32.9%; Pred. No. 5.3;  
 Matches 25; Conservative 8; Mismatches 29; Indels 14; Gaps 5;  
 QY 10 IVODDPRVYRTTTPRGSIPRQSFNKGHAPGPGPGPQQAGARLGAASPFNDL 68  
 DB 133 LREDEKRRRVAGPVGVPLE---GSGRAGPGGFVPSIQ-----GVPESPFRRH 183  
 QY 69 RQI---VNMGFPOMH 81  
 DB 184 EGLDIRGNRGFP-WDI 198  
 RESULT 5  
 AD11\_HUMAN STANDARD; PRT; 769 AA.  
 ID AD11\_HUMAN  
 AC Q75078; Q14808; Q14809; Q14810;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE ADAM 11 precursor (A disintegrin and metalloproteinase domain 11)  
 DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)  
 DE (MDC).  
 GN ADAM11 OR MDC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RX MEDLINE=98359734; PubMed=9693107;  
 RC TISSUE=Brain;  
 RA Sagane K., Ohya Y., Hasegawa Y., Tanaka I.;  
 RA "Metalloproteinase-like, disintegrin-like, cysteine-rich proteins MDC2  
 RT and MDC3: novel human cellular disintegrins highly expressed in the  
 RT brain.";  
 RL Biochem. J. 334:93-98(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=94073190; PubMed=8252040;  
 RA Eml M., Katagiri T., Harada Y., Saito H., Inazawa J., Ito I.,  
 RA Kasumi F., Nakamura Y.;  
 RT "A novel metalloproteinase/disintegrin-like gene at 17q21.3 is  
 RT somatically rearranged in two primary breast cancers.";  
 RL Nat. Genet. 5:151-157(1993).  
 RN [3]  
 RP SEQUENCE OF 106-769 FROM N.A. (ISOFORMS LONG AND SHORT).  
 RC TISSUE=Brain, Breast, Ovary, and Testis;  
 RX MEDLINE=95044425; PubMed=7956356;  
 RA Katagiri T., Harada Y., Eml M., Nakamura Y.;  
 RT "Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron  
 RT organization and alternative splicing.";  
 RL Cytogenet. Cell Gene. 68:39-44(1995).  
 CC -1- FUNCTION: Probable ligand for integrin in the brain. This is a non  
 CC catalytic metalloproteinase-like protein.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Name-Long: Synonyms=MDC-769; Named isoforms=2;  
 CC Name-Short: Synonyms=MDC-524;  
 CC IsoId=O75078-1; Sequence=Displayed;  
 CC IsoId=O75078-2; Sequence=VSP\_005472, VSP\_005473, VSP\_005474,

```

CC VSP_005475;
CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain. Slightly
CC detected or not at all in other tissues.
CC -1- DOMAIN: A conserved motif [AVN(E/D)CD] within the disintegrin-like
CC domain could be involved in the binding to the integrin receptor.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -1- DISEASE: Rearrangements occur in breast and ovarian cancers which
CC involve multiple exons and disrupt the coding region.
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 disintegrin domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB009675; BAA32352.1; -
CC DR EMBL; D17390; BAA04213.1; -
CC DR EMBL; D31872; BAA06670.1; -
CC DR EMBL; D31872; BAA06671.1; -
CC DR PIR; I65967; I65967.
CC DR HSSP; P16619; IFTL.
CC DR MEROPS; M12.976; -.
CC DR Genew; HGNC:189; ADM11.
CC DR MIM; 155120; -.
CC DR GO; GO:0005886; C:plasma membrane; TAS.
CC DR GO; GO:0005178; F:integrin binding; TAS.
CC DR GO; GO:0008237; F:metallopeptidase activity; TAS.
CC DR GO; GO:0007229; F:integrin-mediated signaling pathway; TAS.
CC DR InterPro; IPR006586; ADM_cysteine.
CC DR InterPro; IPR001762; Disintegrin.
CC DR InterPro; IPR006209; EGF_1like.
CC DR InterPro; IPR006025; Pept_M_Zn_BS.
CC DR InterPro; IPR001590; Peptidase_M12B.
CC DR InterPro; IPR002870; Peptidase_M12B_N.
CC DR Pfam; PF00200; disintegrin_1.
CC DR Pfam; PF01562; Pep_M12B_propep; 1.
CC DR Pfam; PF01421; Reptolysin_1.
CC DR PRINTS; PR00289; DISINTEGRIN.
CC DR Prodom; PD000664; Disintegrin; 1.
CC DR SMART; SMO00608; ACR; 1.
CC DR SMART; SMO0050; DISTRN; 1.
CC DR PROSITE; PS00215; ADAM_MERPO; 1.
CC DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC DR PROSITE; PS00214; DISINTEGRIN_2; 1.
CC DR PROSITE; PS00022; EGF_1; 1.
CC DR PROSITE; PS01186; EGF_2; FALSE_NEG.
CC DR PROSITE; PS00026; EGF_3; 1.
CC DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
CC DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
CC DR Signal; glycoprotein; Transmembrane; EGF-like domain;
CC Alternative splicing.
CC KW SIGNAL; 1 23
CC FT PROPEP; 24 225
CC FT CHAIN; 226 769
CC FT DOMAIN; 226 734
CC FT TRANSMEM; 735 755
CC FT DOMAIN; 756 769
CC FT DOMAIN; 226 438
CC FT DOMAIN; 444 531
CC FT DOMAIN; 532 676
CC FT DOMAIN; 677 709
CC FT DISULFID; 349 433
CC FT DISULFID; 503 516
CC FT DISULFID; 677 692
CC FT DISULFID; 686 698
CC FT DISULFID; 700 709
CC FT CARBOHYD; 96
CC N-LINKED (GLCNAC...) (POTENTIAL).

```

```

FT CARBOHYD 605 605 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPPLIC 1 99 Missing (in isoform Short).
FT VARSPPLIC 100 104 /FtId=VSP_005472.
FT VARSPPLIC 100 104 /FtId=VSP_005473.
FT VARSPPLIC 595 623 DVLGFLICVINSIGAPRGDVGIVSVT -> PQQGRAYV
FT VARSPPLIC 624 769 LPLICHLWSSSARPGGRHQ (in isoform Short).
FT VARSPPLIC 624 769 Missing (in isoform Short).
FT CONFLICT 106 106 /FtId=VSP_005475.
FT CONFLICT 325 325 Q -> H (IN REF. 1).
FT SEQUENCE 769 AA; 83409 MW; 59B9C467B6DDFB6E CRC64;
SO SEQUENCE 769 AA; 83409 MW; 59B9C467B6DDFB6E CRC64;

Query Match 11.4%; Score 72; DB 1; Length 769;
Best Local Similarity 27.7%; Pred. No. 20;
Matches 23; Conservative 10; Mismatches 40; Indels 10; Gaps 2;

QY 8 GRVQDDPRVYTTTPQPR-----GSIPQSPFNRGHAPPGAPGPGQAGARLGA 60
DB 34 GGLPQLGGPAPPEVTEPRRLVRESGGEVRKQDTRVRQSPGAPPVHLAQVGFVIAF 93
QY 61 QSPFN---DLRQLYVMGFPQWH 80
DB 94 NSNFTLDLNLHQLLSQVERH 116

RESULT 6
P1A0_RABIT STANDARD; PRT; 490 AA.
ID P1A0_RABIT
AC 019045;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN F10.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97256311; PubMed=9101642;
RA Pendurthi U.R., Anderson K.D., James H.L.;
RT "Characterization of a full-length cDNA for rabbit factor X.";
RL Thromb. Res. 85:503-514(1997).
CC -1- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC converts prothrombin to thrombin in the presence of factor Va,
CC calcium and phospholipid during blood clotting.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC Arg-|-Ile bonds in prothrombin to form thrombin.
CC -1- SUBUNIT: The two chains are formed from a single-chain precursor
CC by the excision of two Arg residues and are held together by 1 or
CC more disulfide bonds.
CC -1- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC glutamate residues allows the modified protein to bind calcium (By
CC similarity).
CC -1- PTM: N- and O-glycosylated (By similarity).
CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)
CC (BY SIMILARITY).
CC -1- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC another site, beyond the GLA domain.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

CC or send an email to license@isb-sib.ch).

DR EMBL; AF003200; AAB62542.1; -.

DR HSPB; P00742; HSCG.

DR MEROPS; S01.216; -.

DR InterPro; IPRO00152; Asx\_hydroxyl\_S.

DR InterPro; IPRO00903; Cys\_ser\_trypsin.

DR InterPro; IPRO00742; EGF\_2.

DR InterPro; IPRO01861; EGF\_Ca.

DR InterPro; IPRO01438; EGF\_T1.

DR InterPro; IPRO06209; EGF\_like.

DR InterPro; IPRO02383; GLA\_blood.

DR InterPro; IPRO01254; peptidase\_S1.

DR InterPro; IPRO01314; peptidase\_S1A.

DR InterPro; IPRO00294; VitK\_dep\_GLA.

DR Pfam; PF00008; EGF\_2.

DR Pfam; PF00584; gla\_1.

DR Pfam; PF00089; trypsin\_1.

DR PRINTS; PR00722; CHYOTRYPsin.

DR PRINTS; PR00010; EGFBL00D.

DR SMART; SM000179; EGF\_CA; 1.

DR SMART; SM00069; GLA; 1.

DR SMART; SM00020; TRYP\_SRC; 1.

DR PROSITE; PS00010; ASX\_HYDROXYL; 1.

DR PROSITE; PS00022; EGF\_1; 1.

DR PROSITE; PS01186; EGF\_2; 2.

DR PROSITE; PS50026; EGF\_3; 1.

DR PROSITE; PS01187; EGF\_CA; 1.

DR PROSITE; PS00011; GLUT\_CARBOXYLATION; 1.

DR PROSITE; PS50240; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; 1.

DR PROSITE; PS00133; TRYPSIN\_SER; 1.

KM Gamma-carboxypeptidase; Serine protease; Plasma; blood coagulation;

KM Gamma-carboxypeptidase; Hydroxylase; Hydroxylation; Calcium-binding; Vitamin K;

KM Signal; Zymogen; EGF-like domain. Repeat.

FT SIGNAL 1 20 POTENTIAL.

FT PROPEP 21 40 BY SIMILARITY.

FT CHAIN 41 180 FACTOR X LIGHT CHAIN.

FT CHAIN 184 490 FACTOR X HEAVY CHAIN.

FT PROPEP 184 232 ACTIVATION PEPTIDE.

FT CHAIN 233 490 ACTIVATED FACTOR XA, HEAVY CHAIN.

FT DOMAIN 86 122 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 125 165 EGF-LIKE 2.

FT DOMAIN 233 490 SERINE PROTEASE.

FT MOD\_RES 46 46 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 47 47 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 54 54 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 56 56 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 59 59 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 65 65 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 72 72 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 75 75 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 79 79 GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD\_RES 103 103 HYDROXYLATION (BY SIMILARITY).

FT ACT\_SITE 274 274 CHARGE RELAY SYSTEM.

FT ACT\_SITE 320 320 CHARGE RELAY SYSTEM.

FT ACT\_SITE 417 417 CHARGE RELAY SYSTEM.

FT DISULFID 90 101 BY SIMILARITY.

FT DISULFID 95 110 BY SIMILARITY.

FT DISULFID 112 121 BY SIMILARITY.

FT DISULFID 126 140 BY SIMILARITY.

FT DISULFID 136 149 BY SIMILARITY.

FT DISULFID 151 164 BY SIMILARITY.

FT DISULFID 172 340 INTERCHAIN (BY SIMILARITY).

FT DISULFID 239 244 BY SIMILARITY.

FT DISULFID 259 275 BY SIMILARITY.

FT DISULFID 388 402 BY SIMILARITY.

FT DISULFID 413 441 BY SIMILARITY.

FT CARBOHYD 61 61 N-LINKED (GLUCNAc. . .) (POTENTIAL).

FT CARBOHYD 187 187 N-LINKED (GLUCNAc. . .) (POTENTIAL).

FT CARBOHYD 205 205 N-LINKED (GLUCNAc. . .) (POTENTIAL).

SQ SEQUENCE 490 AA; 53965 MW; 3A39FA85AFA2A6D11 CRC64;

Query Match 11.3%; Score 71.5; DB 1; Length 490;

Best Local Similarity 32.8%; Pred. No. 14;

Matches 22; Conservative 9; Mismatches 31; Indels 5; Gaps 2;

QY 6 ANGEIVODDPRVRRTPPPGSGIP---RQSFNRRGAGPPGPGPPQQAAGARLGAQA 61

Db 152 ASGYTLGNNGKSCISTELFGGKVTLLGWRSPANNSSEGPPEARGFQDDG-VLTATE 210

QY 62 SPFDLN 68

Db 211 NPFNLD 217

RESULT 7

MC22\_SALTY STANDARD; PRT; 553 AA.

ID MC22\_SALTY

AC P02941;

DT 21-JUL-1986 (Rel. 01, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Methyl-accepting chemotaxis protein II (MCP-II) (Aspartate

DE chemoreceptor protein).

GN TAR OR CHEM OR STM1919.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI\_Taxid=602;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83197387; PubMed=6302843;

RA Russo A.F., Koshland D.E. Jr.;

RT "Separation of signal transduction and adaptation functions of the

RL aspartate receptor in bacterial sensing.";

RL Science 220:1016-1020 (1983).

RN [2]

RP REVISIONS TO 236 AND INSERTION OF ALA-498.

RC STRAIN=LT2;

RA Stock A.;

RN Thesis (1986), University of Berkeley, U.S.A.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=LT2 / SGSC1412 / ATCC 700720;

RX MEDLINE=21534948; PubMed=11677609;

RA McClelland W., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvanev E.,

RA Ryan E., Sun H., Florea L., Miller W., Stonking T., Nhan M.,

RA Waterston R., Wilson R.K.;

RT "Complete genome sequence of Salmonella enterica serovar Typhimurium

RT LT2.";

RL Nature 413:852-856 (2001).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 26-188.

RX MEDLINE=92073901; PubMed=1660187;

RA Wilburn M.V., Prive G.G., Milligan D.L., Scott W.G., Yeh J.,

RA Tancarik J., Koshland D.E. Jr., Kim S.H.;

RT "Three-dimensional structures of the ligand-binding domain of the

```

RT bacterial aspartate receptor with and without a ligand.",
RL Science 254:1342-1347(1991).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 36-180.
RX MEDLINE=93252857; PubMed=8486661;
RA Yeh J.I., Biemann H.-P., Pandit J., Koshland D.E., Kim S.-H.;
RT "High-resolution structures of the ligand binding domain of the wild-
RT type bacterial aspartate receptor.";
RT J. Mol. Biol. 262:186-201(1996).
CC -1- FUNCTION: RECEPTOR FOR THE ATTRACTANT L-ASPARTATE AND RELATED
CC AMINO AND DICARBOXYLIC ACIDS. TAR MEDIATES TAXIS AWAY FROM THE
CC REPELLENTS COBALT AND NICKEL. UNLIKE E. COLI TAR IT DOES NOT
CC MEDIATES MALTOSE TAXIS.
CC -1- FUNCTION: CHEMOTACTIC-SIGNAL TRANSDUCERS RESPOND TO CHANGES IN THE
CC CONCENTRATION OF ATTRACTANTS AND REPELLENTS IN THE ENVIRONMENT",
CC TRANSDUCE A SIGNAL FROM THE OUTSIDE TO THE INSIDE OF THE CELL, AND
CC FACILITATE SENSORY ADAPTATION THROUGH THE VARIATION OF THE LEVEL
CC OF METHYLATION. ATTRACTANTS INCREASE THE LEVEL OF METHYLATION
CC WHILE REPELLENTS DECREASE THE LEVEL OF METHYLATION, THE METHYL
CC GROUPS ARE ADDED BY THE METHYLTRANSFERASE CHER AND REMOVED BY THE
CC METHYLTRANSFERASE CHEB.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -1- SIMILARITY: Contains 1 HAMP domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, J01809; ; NOT ANNOTATED_CDS.
DR EMBL, AE008785; AAL20835.1; -.
DR PDB, 1WAS; 15-JAN-95.
DR PDB, 1WAT; 15-JAN-95.
DR PDB, 2LIG; 15-SEP-95.
DR PDB, 1LIH; 15-SEP-95.
DR PDB, 1VLS; 21-APR-97.
DR PDB, 1VLT; 15-MAY-97.
DR PDB, 1JMW; 22-AUG-01.
DR StyGene; SG10387; tax.
DR InterPro; IPR004091; AspME chemerecept.
DR InterPro; IPR004089; CmtTaxis_transd.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR004090; Me_chemotaxis.
DR InterPro; IPR003122; Tatr.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR Pfam; PF02203; Tatr; 1.
DR PRINTS; PR00260; CHEMTRNSDCR.
DR SMART; SMO0304; HAMP; 1.
DR SMART; SMO0283; MA; 1.
DR SMART; SMO0319; Tatr; 1.
DR PROSITE; PS00538; CHEMOTAXIS_TRANSDUC_1; 1.
DR PROSITE; PS00111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 1.
KW Chemotaxis; Transducer; Transmembrane; Methylation; Periplasmic;
KW Inner membrane; 3D-structure; Complete proteome.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 33 POTENTIAL.
FT PERIPLASMIC 34 190 PERIPLASMIC (POTENTIAL).

```

| FT                    | TRANSMEM                                                                                                                 | 191                                                        | 211          | POTENTIAL.                                                                                                |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|--------------|-----------------------------------------------------------------------------------------------------------|
| FT                    | DOMAIN                                                                                                                   | 212                                                        | 553          | CYTOPLASMIC (POTENTIAL).                                                                                  |
| FT                    | DOMAIN                                                                                                                   | 214                                                        | 266          | HAMP.                                                                                                     |
| FT                    | DOMAIN                                                                                                                   | 271                                                        | 500          | METHYL-ACCEPTING TRANSUCER.                                                                               |
| FT                    | SITE                                                                                                                     | 64                                                         | 73           | THE 3 ARG MAY FORM A POSITIVELY CHARGED POCKET, WHICH BINDS THE ALPH-CARBOXYL GROUP OF THE ATTRACTANT AA. |
| FT                    | MOD_RES                                                                                                                  | 295                                                        | 295          | DEAMINATION AND METHYLATION.                                                                              |
| FT                    | MOD_RES                                                                                                                  | 302                                                        | 302          | DEAMINATION AND METHYLATION.                                                                              |
| FT                    | MOD_RES                                                                                                                  | 309                                                        | 309          | METHYLATION.                                                                                              |
| FT                    | MOD_RES                                                                                                                  | 491                                                        | 491          | E -> G (IN REF. 1).                                                                                       |
| FT                    | CONFLICT                                                                                                                 | 246                                                        | 246          | T -> A (IN REF. 1).                                                                                       |
| FT                    | CONFLICT                                                                                                                 | 451                                                        | 451          |                                                                                                           |
| FT                    | HELI                                                                                                                     | 36                                                         | 38           |                                                                                                           |
| FT                    | HELI                                                                                                                     | 39                                                         | 73           |                                                                                                           |
| FT                    | HELI                                                                                                                     | 76                                                         | 77           |                                                                                                           |
| FT                    | TURN                                                                                                                     | 82                                                         | 83           |                                                                                                           |
| FT                    | HELI                                                                                                                     | 89                                                         | 109          |                                                                                                           |
| FT                    | HELI                                                                                                                     | 114                                                        | 116          |                                                                                                           |
| FT                    | HELI                                                                                                                     | 117                                                        | 142          |                                                                                                           |
| FT                    | TURN                                                                                                                     | 143                                                        | 144          |                                                                                                           |
| FT                    | HELI                                                                                                                     | 146                                                        | 151          |                                                                                                           |
| FT                    | HELI                                                                                                                     | 154                                                        | 177          |                                                                                                           |
| FT                    | TURN                                                                                                                     | 178                                                        | 179          |                                                                                                           |
| FT                    | SEQUENCE                                                                                                                 | 553 AA;                                                    | 59614 MW;    | SCCE8E707452760D CRC64;                                                                                   |
| Query Match           |                                                                                                                          | 11.3%;                                                     | Score 71.5;  | DB 1; Length 553;                                                                                         |
| Best Local Similarity |                                                                                                                          | 28.2%;                                                     | Pred. No.16; |                                                                                                           |
| Matches               | 22;                                                                                                                      | Conservative                                               | 15;          | Mismatches 28; Indels 13; Gaps 3;                                                                         |
| QY                    | 50                                                                                                                       | QQAAGARIGAAQSPENDINQVLVNG----QPMHIGNHAHEPVTISLLEFILMNL--G  | 102          |                                                                                                           |
| DB                    | 158                                                                                                                      | QNALEBAAGNARVASENLYRQTFQSDAHYRFAQWOLG-----VLAVALVLIILMVWFG | 211          |                                                                                                           |
| QY                    | 103                                                                                                                      | VRGLLVGVNYLVSHLSQ                                          | 120          |                                                                                                           |
| DB                    | 212                                                                                                                      | IRHALLNPLARVTHIRE                                          | 229          |                                                                                                           |
| RESULT 8              |                                                                                                                          |                                                            |              |                                                                                                           |
| 34KD MYCPA            |                                                                                                                          |                                                            |              |                                                                                                           |
| ID                    | 34KD MYCPA                                                                                                               | STANDARD;                                                  | PRT;         | 298 AA.                                                                                                   |
| AC                    | 004959;                                                                                                                  |                                                            |              |                                                                                                           |
| DT                    | 01-JUN-1994 (Rel. 29, Created)                                                                                           |                                                            |              |                                                                                                           |
| DT                    | 01-JUN-1994 (Rel. 29, Last sequence update)                                                                              |                                                            |              |                                                                                                           |
| DT                    | 30-MAY-2000 (Rel. 39, Last annotation update)                                                                            |                                                            |              |                                                                                                           |
| DE                    | 34 kDa antigenic protein.                                                                                                |                                                            |              |                                                                                                           |
| OS                    | Mycobacterium paratuberculosis.                                                                                          |                                                            |              |                                                                                                           |
| OC                    | Bacteria; Actinobacteria; Actinomycetales;                                                                               |                                                            |              |                                                                                                           |
| OC                    | Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.                                                                     |                                                            |              |                                                                                                           |
| OX                    | NCBI_TaxID=1770;                                                                                                         |                                                            |              |                                                                                                           |
| RN                    | [1]                                                                                                                      |                                                            |              |                                                                                                           |
| RP                    | SEQUENCE FROM N.A.                                                                                                       |                                                            |              |                                                                                                           |
| RX                    | MEDLINE=93328703; PubMed=8335649;                                                                                        |                                                            |              |                                                                                                           |
| RA                    | Gillet P., de Kessel M., Machtelinckx L., Coene M., Coclico C.;                                                          |                                                            |              |                                                                                                           |
| RA                    | Isolation and sequencing of the gene coding for an antigenic 34-                                                         |                                                            |              |                                                                                                           |
| RT                    | kilodalton protein of Mycobacterium paratuberculosis."                                                                   |                                                            |              |                                                                                                           |
| RL                    | J. Bacteriol. 175:4930-4935(1993).                                                                                       |                                                            |              |                                                                                                           |
| CC                    | -1- SUBCELLULAR LOCATION: Integral membrane protein.                                                                     |                                                            |              |                                                                                                           |
| CC                    | -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV0954.                                                                        |                                                            |              |                                                                                                           |
| CC                    | -----                                                                                                                    |                                                            |              |                                                                                                           |
| CC                    | This SWISS-PROT entry is copyright. It is produced through a collaboration                                               |                                                            |              |                                                                                                           |
| CC                    | between the Swiss Institute of Bioinformatics and the EMBL collaboration                                                 |                                                            |              |                                                                                                           |
| CC                    | the European Bioinformatics Institute. There are no restrictions on its                                                  |                                                            |              |                                                                                                           |
| CC                    | use by non-profit institutions as long as its content is in no way                                                       |                                                            |              |                                                                                                           |
| CC                    | modified and this statement is not removed. Usage by and for commercial                                                  |                                                            |              |                                                                                                           |
| CC                    | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |                                                            |              |                                                                                                           |
| CC                    | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).                                        |                                                            |              |                                                                                                           |
| CC                    | -----                                                                                                                    |                                                            |              |                                                                                                           |
| DR                    | EMBL; X68102; CAA48221.1; -.                                                                                             |                                                            |              |                                                                                                           |
| DR                    | PIR; A40615; A40615.                                                                                                     |                                                            |              |                                                                                                           |
| KW                    | Antigen; Transmembrane.                                                                                                  |                                                            |              | POTENTIAL.                                                                                                |
| FT                    | TRANSMEM                                                                                                                 | 42                                                         | 62           |                                                                                                           |

FT TRANSMEM 77 97 POTENTIAL.  
 FT TRANSMEM 102 122 POTENTIAL.  
 FT TRANSMEM 134 154 POTENTIAL.  
 SQ SEQUENCE 298 AA; 29565 MW; B2E57AA2077F0D3D CRC64;

Query Match 11.2%; Score 71; DB 1; Length 298;  
 Best Local Similarity 27.5%; Pred. No. 9.4;  
 Matches 30; Conservative 13; Mismatches 34; Indels 32; Gaps 4;

DB 43 PGGPG-PROCGAGARLGAAGSPFN-----LNRLQVNMGPFWML----- 81  
 5 PGGPGVPPAOSGTAVGATPSPAKDDGSKPLPLYNIAVVALGPAAYLNGPFTTGA 64

QY 82 -----GNNAVEPTSTLLFLMLGVGRL-----LVGLVYVSHS 119  
 65 DLPGPGIGRAGDAGTAVVALLAALLAGLLPKAXYGVAVVAVALA 113

DB 65 DLPGPGIGRAGDAGTAVVALLAALLAGLLPKAXYGVAVVAVALA 113

RESULT 9  
 GAST\_HUMAN STANDARD; PRT; 345 AA.  
 ID GAST\_HUMAN  
 AC P54826;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Growth-arrest-specific protein 1 precursor (GAS-1).  
 GN GAST.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RC TISSUE=Liver;  
 RX MEDLINE=94173926; PubMed=8127893;  
 RA del Sal G., Collavin L., Ruaro M.E., Edom P., Saccone S., Valle G.D.,  
 RA Schneider C.;  
 RT "Structure, function, and chromosome mapping of the  
 RT growth-suppressing human homologue of the murine gas1 gene";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1848-1852(1994).  
 CC -1- FUNCTION: Specific growth arrest protein involved in growth  
 CC suppression. Blocks entry to S phase. Prevents cycling of  
 CC normal and transformed cells.  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By  
 CC similarity).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L13698; AAA72368.1; -;  
 DR PIR; A53138; A53138.  
 DR Genew; HGNC:4165; GAST.  
 DR MIM; 139185; -;  
 DR GO; GO:0046658; C:extrinsic to plasma membrane, GPI-anchored; ISS.  
 DR GO; GO:0007050; P:cell cycle arrest; TAS.  
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 DR GO; GO:0045749; P:negative regulation of S phase of mitotic c...; NAS.  
 KM Growth arrest; GPI-anchor; Signal; Glycoprotein; Lipoprotein.  
 FT SIGNAL 1 39  
 FT CHAIN 40 318  
 FT PROPEP 319 345  
 FT DOWAIN 87 285  
 FT DOWAIN 285 288  
 FT DOWAIN 319 322  
 FT DOWAIN 337 341  
 FT LIPID 318 318  
 FT CARBOHYD 117 117  
 FT SEQUENCE 345 AA; 35721 MW; 2AADS0FLD3632F9D CRC64;

Query Match 11.2%; Score 71; DB 1; Length 345;  
 Best Local Similarity 29.3%; Pred. No. 11;  
 Matches 27; Conservative 4; Mismatches 33; Indels 28; Gaps 2;

DB 6 ANGEIVDDDDPRVYTTTPPGSGIPROSFNKGAPGPGPGPQAGARLGAAGSPFN 65  
 278 AGSGQPLDDDDGVHPHPPSGGAA-----SGRGDLPYFPGRSSGGGRLAP----- 326

QY 66 DLNRLQVNMGPFWMLGNNAVEPTSTLLFL 97  
 327 -----RGAWTPILASILLILL 341

DB 327 -----RGAWTPILASILLILL 341

RESULT 10  
 CTF8\_HUMAN STANDARD; PRT; 478 AA.  
 ID CTF8\_HUMAN  
 AC Q96D72; Q9UP6;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Protein C20orf158.  
 GN C20ORF158.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Bloecher H., Boecher M., Brandt P., Mewes H.-W., Gassenhuber J.,  
 RA Wiemann S.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.,  
 RA Jones M., Stavaris G., Almeida J.P., Babbage A.K., Baggeley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Levasainio M.H., Leverha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McCormackie L.J., McElay K., Murray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.,  
 RA Wilmberg L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20";  
 RU Nature 414:865-871 (2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Matrusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton W., Soares W.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA White J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL133063; CAB61387.1; ALT\_INIT.  
 DR EMBL; AL035669; CAC12754.1; ALT\_INIT.  
 DR EMBL; BC012757; AAH12757.1; -.  
 DR Genew; HANC:16151; C200rf158. PRO-RICH.  
 FT DOMAIN 7 272  
 FT DOMAIN 346 452 ARG-RICH.  
 FT CONFLICT 253 253 A -> P (IN REF. 3).  
 SQ SEQUENCE 478 AA; 53940 MW; EEC03CFD70BA3F00 CRC64;

Query Match 11.2%; Score 71; DB 1; Length 478;  
 Best Local Similarity 35.4%; Pred. No. 15;  
 Matches 23; Conservative 4; Mismatches 24; Indels 14; Gaps 4;

QY 10 IVDDDDPRVATTQPP-----RGSIPRPSFNRHGAPPGP--GPRQQAARLGAAG 61  
 DB 24 IASNDGR-----GPPPARFAGAKGPIP--SLFSGQSPPPYDSKSPSYLGGPPGVAP 77  
 QY 62 SPND 66  
 DB 78 SOFE 82

RESULT 11  
 ID CBL2\_MOUSE STANDARD; PRT; 481 AA.  
 AC Q8K3M5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cdk5 and Abl enzyme substrate 2 (Interactor with cdk3 2) (IK3-2).  
 GN CABL852.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver, and Spleen;  
 RX MEDLINE=21952370; PubMed=11955625;  
 RA Sato H., Nishimoto I., Matsuo K.,  
 RT "Ik3-2, a relative to ik3-1/cables, is associated with cdk3, cdk5, and  
 RT c-abl-2";  
 RL Biochim. Biophys. Acta 1574:157-163(2002).  
 RL FUNCTION: Unknown. Probably involved in G1-S cell cycle  
 RL transition.  
 CC -1- SUBUNIT: Binds to CDK3, CDK5 and c-Abl. The C-terminal cyclin-box-  
 CC -1- TISSUE SPECIFICITY: Widely expressed.  
 CC -1- SIMILARITY: Belongs to the cyclin family

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AY049712; AAL12171.1; -.  
 DR MGD; MGI:218235; Cables2.  
 DR GO; GO:0005515; F:protein binding; IPT.  
 DR InterPro; IPR006671; Cyclin\_N.  
 DR Pfam; PF00134; cyclin; 1.  
 KW Cell cycle; Cell division; Cyclin.  
 FT DOMAIN 7 11 POLY-ALA.  
 FT DOMAIN 15 113 PRO-RICH.  
 SQ SEQUENCE 481 AA; 52710 MW; 00DALT706578B1B2 CRC64;

Query Match 11.2%; Score 70.5; DB 1; Length 481;  
 Best Local Similarity 35.9%; Pred. No. 17;  
 Matches 23; Conservative 5; Mismatches 21; Indels 15; Gaps 4;

QY 25 PRGSPRPSFNRHGAPPGP--GPR-RQQAARLGAAGSPNDLRQVNMGFQWH 80  
 DB 269 PRPSTPR-----APPSRHRKVPPTKSTPAGTELGSDGDAVEYNPLLD--DPQWP 317  
 QY 81 LGNH 84  
 DB 318 CGKH 321

RESULT 12  
 ID S24A\_ARATH STANDARD; PRT; 1054 AA.  
 AC Q9SFU0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative protein transport protein Sec24-like At3g07100.  
 GN AT3G07100 OR T1B9.25.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016720; PubMed=11130713;  
 RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unsel M.,  
 RA Pattmann B., Valle G., Bloeker H., Perez-Alonso M., Obermair B.,  
 RA Deisner M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,  
 RA De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P.,  
 RA Winkler P., Catolico L., Weissenbach J., Saurin W., Quetier F.,  
 RA Schaefer M., Muller-Auer S., Gabel C., Fuchs M., Benes V.,  
 RA Mumbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,  
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,  
 RA Verzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,  
 RA Conrad A., Hornischer K., Kauer G., Loehmet T.H., Nordstiek G.,  
 RA Reichelt J., Scharte M., Schoen O., Bargues M., Terol J., Clement J.,  
 RA Navarro P., Collado C., Perez-Perez A., Ockenwelder B., Duchemin D.,  
 RA Cooke R., laudie M., Berger-Liauro C., Purnelle B., Masny D.,  
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
 RA Morfot A., Argiriou A., Flores M., Liguori R., Vitale D.,  
 RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Vewes H.-M.,  
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Rooney T., Rizzo M., Walte A., Ulteback T., Fujii C.Y., Shea T.P.,  
 RA Cerasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,  
 RA Pal G., Miltischer J., Sellers P., Gill J.E., Feldlyum T.V.,  
 RA Preuss D., Iln X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,  
 RA Fraser C.M., Kaneke T., Nakamura Y., Saito S., Kato T., Asamizu F.

RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shino S., Takenuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 thaliana.";  
 RL Mature 408,820-822(2000).  
 CC -1- FUNCTION: Component of the COP1 coat, that covers ER-derived  
 CC vesicles involved in transport from the endoplasmic reticulum to  
 CC the Golgi apparatus. COP1 is composed of at least five proteins:  
 CC the SEC23/24 complex, the SEC13/31 complex, and the protein SARL.  
 CC Acts in the cytoplasm to promote the transport of secretory,  
 CC plasma membrane, and vacuolar proteins from the endoplasmic  
 CC reticulum to the Golgi complex (By similarity).  
 CC -1- SIMILARITY: Belongs to the SEC23/SEC24 family. SEC24 subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AC012395; AAF20236.1; -  
 DR InterPro; IPR007123; Gelsolin.  
 DR InterPro; IPR006900; Sec23\_helical.  
 DR InterPro; IPR006895; Sec23\_trunk.  
 DR InterPro; IPR006895; ZF-Sec23\_Sec24.  
 DR Pfam; PF00626; Gelsolin; 1.  
 DR Pfam; PF04815; Sec23\_helical; 1.  
 DR Pfam; PF04811; Sec23\_trunk; 1.  
 DR Pfam; PF04810; ZF-Sec23\_Sec24; 1.  
 DR Hypothetical protein; Transport; Protein transport; Golgi stack;  
 KW Endoplasmic reticulum; Multigene family.  
 FT DOMAIN 369 394 ZINC FINGER-LIKE.  
 FT DOMAIN 89 92 POLY-PRO.  
 FT DOMAIN 128 131 POLY-PRO.  
 SQ SEQUENCE 1054 AA; 115662 MW; 13FAB74147EAB9C8 CRC64;  
 QY Query Match 11.1%; Score 70; DB 1; Length 1054;  
 Best Local Similarity 30.7%; Pred. No. 44;  
 Matches 23; Conservative 6; Mismatches 32; Indels 14; Gaps 3;  
 QY 16 PRVTTTTPPGSGIPROSFNFGHAPPGGPPGQOAGARLGAAGSPFNDLNRQLVNG 75  
 DB 116 PPPPTTNPFGGPPPTTL--AGHLSPMSLRPQGPAPVAVMGPPP-----GTTSG 165  
 QY 76 FPGWHLNHAVERPT 90  
 DB 166 LP---GANAYPPAT 176  
 RESULT 13  
 AANT\_HDVNA STANDARD; PRT; 214 AA.  
 ID AANT\_HDVNA  
 AC P25880;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-MAY-1992 (Rel. 22, Last annotation update)  
 DE Delta antigen.  
 OS Hepatitis delta virus (isolate Nauru) (HDV).  
 OC Viruses; Deltaavirus.  
 OC NCBI\_TaxID=10426;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91020976; PubMed=2219700;  
 RA Chao Y.C., Chang M.F., Gust I., Lai M.M.C.;  
 RT "Sequence conservation and divergence of hepatitis delta virus RNA.";  
 RL Virology 178:384-392(1990).  
 CC -1- SUBCELLULAR LOCATION: Nucleus.  
 CC -1- PTM: Phosphorylated.  
 CC -1- DISEASE: BINDS SPECIFICALLY TO ANTISERA FROM PATIENTS WITH CHRONIC  
 CC HEPATITIS DELTA VIRAL INFECTIONS.

CC -1- MISCELLANEOUS: Delta antigen binds specifically to HDV RNA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M58629; AAB59753.1; -  
 DR PIR; A36212; SAVDND.  
 DR InterPro; IPR002506; HDV\_ag.  
 DR Pfam; PF01517; HDV\_ag; 1.  
 DR Pfam; PD002867; HDV\_ag; 1.  
 KW Antigen; RNA-binding; Nuclear protein; Phosphorylation.  
 SQ SEQUENCE 214 AA; 24180 MW; 91F2D28B55E801D CRC64;  
 QY Query Match 10.9%; Score 69; DB 1; Length 214;  
 Best Local Similarity 29.3%; Pred. No. 10;  
 Matches 22; Conservative 8; Mismatches 33; Indels 12; Gaps 4;  
 QY 10 IVDDDFVFRITTPPGSGIPROSFNFGHAPPGGPPGQOAGARLGAAGSPFNDLNR 69  
 DB 133 LTBEDRRERVRVAGPPGCV--NPLEGSGRGAFGGFPVPMQ-----GVESPTRTGE 184  
 QY 70 QL---VNMGPQWHL 81  
 DB 185 GLDVRGDRGF-WDI 198  
 RESULT 14  
 IPF1\_HUMAN STANDARD; PRT; 283 AA.  
 ID IPF1\_HUMAN  
 AC P52945; O60594;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Insulin promoter factor-1 (IPF-1) (Pancreas/duodenum homeobox-1) (PDX-  
 DE 1) (Islet/duodenum homeobox-1) (IDX-1) (Somatostatin transactivating  
 DE factor-1) (STF-1) (Insulin upstream factor-1) (IUF-1) (Glucose-  
 DE sensitive factor) (GSF).  
 GN IPF1 OR PDX1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96070447; PubMed=7590740;  
 RA Stoffel M., Stein R., Wright C.V., Espinosa R. III, le Beau M.M.,  
 RA Bell G.I.;  
 RT "Localization of human homeodomain transcription factor insulin  
 RT promoter factor 1 (IPF1) to chromosome band 13q12.1.";  
 RL Genomics 28:125-126(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pancreatic islets;  
 RX MEDLINE=96220081; PubMed=8635654;  
 RA Inoue H., Riggs A.C., Tanikawa Y., Ueda K., Kuwano A., Liu L.,  
 RA Donis-Keller H., Pernutt M.A.;  
 RT "Isolation, characterization, and chromosomal mapping of the human  
 RT insulin promoter factor 1 (IPF-1) gene.";  
 RL Diabetes 45:789-794(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pancreatic islets;  
 RA Hiroshi I.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pancreatic islets;  
 RA Marshak S., Totary H., Cerasi E., Melloul D.;



Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
[5] SEQUENCE FROM N.A.  
RA Hata M., Linher T.H., Paz V.P., Wang X., Iwaseki N., Bell G.I.;  
Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
[6] SEQUENCE FROM N.A.  
RA TISSUE-Pancreatic Islets;  
RX MEDLINE=97400587; PubMed=9252422;  
RA Macfarlane W.M., Smith S.B., James R.F., Clifton A.D., Doza Y.N.,  
Cohen P., Docherty K.;  
RA "The p38/activating kinase mitogen-activated protein kinase cascade  
mediates the activation of the transcription factor insulin upstream  
factor 1 and insulin gene transcription by high glucose in pancreatic  
beta-cells."  
J. Biol. Chem. 272:20936-20944(1997).  
[7] VARIANTS MODY4 ARG-18, ASN-76 AND HIS-197.  
RA MEDLINE=20015025; PubMed=10545530;  
RA Macfarlane W.M., Frayling T.M., Ellard S., Evans J.C., Allen L.I.,  
Bulman M.P., Ayres S., Shepherd M., Clark P., Millward A., Demaine A.,  
Wilkin T., Docherty K., Hattersley A.T.;  
RA "Missense mutations in the insulin promoter factor-1 gene predispose  
to type 2 diabetes."  
J. Clin. Invest. 104:R33-R39(1999).  
[8] VARIANTS MODY4 LEU-59; ASN-76 AND PRO-243 INS.  
RX MEDLINE=20015026; PubMed=10545531;  
RA Hant E.H., Scofield D.A., Chevre J.-C., Durand E., Stanojevic V.,  
Dina C., Habener J.F., Froguel P.;  
RA "Defective mutations in the insulin promoter factor-1 (IPF-1) gene in  
late-onset type 2 diabetes mellitus."  
J. Clin. Invest. 104:R41-R48(1999).  
[9] FUNCTION: Activates insulin, somatostatin, glucokinase, islet  
amyloid polypeptide and glucose transporter type 2 gene  
transcription. Particularly involved in glucose-dependent  
regulation of insulin gene transcription. Binds preferentially the  
DNA motif 5'-[CT]TAAT[TC]-3'. During development, specifies the  
early pancreatic epithelium, permitting its proliferation,  
branching and subsequent differentiation. At adult stage, required  
for maintaining the hormone-producing phenotype of the beta-cell.  
[10] SUBUNIT: Interacts with the basic helix-loop-helix domains of  
TCF3(E47) and NEUROD1 and with HMG-I(Y) (By similarity).  
[11] SUBUNIT: Interacts with the basic helix-loop-helix domains of  
TCF3(E47) and NEUROD1 and with HMG-I(Y) (By similarity).  
[12] TISSUE SPECIFICITY: Duodenum and pancreas (Langerhans islet beta  
cells and small subsets of endocrine non-beta-cells, at low levels  
in acinar cells).  
[13] DOMAIN: The Amp-type hexapeptide mediates heterodimerization of  
IPF1 with PDX on a regulatory element of the somatostatin promoter  
(By similarity).  
[14] DOMAIN: The homeodomain, which contains the nuclear localization  
signal, not only mediates DNA-binding, but also acts as a protein-  
protein interaction domain for TCF3(E47), NEUROD1 and HMG-I(Y) (By  
similarity).  
[15] PPM: PHOSPHORYLATED BY THE SAPK2 PATHWAY AT HIGH GLUCOSE  
CONCENTRATION.  
[16] DISEASE: Defects in IPF1 are a cause of pancreatic agenesis  
(MIM:260370). This autosomal recessive disorder is characterized  
by absence or hypoplasia of pancreas, leading to early-onset  
insulin-dependent diabetes mellitus. This was found in a  
frameshift mutation that produces a truncated protein and results  
in a second initiation that produces a second protein that act as  
a dominant negative mutant.  
[17] DISEASE: Defects in IPF1 are a cause of susceptibility to  
noninsulin-dependent diabetes mellitus (NIDDM) (MIM:125853); also  
known as diabetes mellitus type II.  
[18] DISEASE: Defects in IPF1 are the cause of maturity onset diabetes  
of the young type IV (MODY4) (MIM:606392); also symbolized MODY-  
IV. MODY4 is a form of non-insulin-dependent diabetes mellitus  
(NIDDM) characterized by an autosomal dominant mode of  
inheritance, age of onset of 25 years or younger and a primary  
defect in insulin secretion.

IPF1/XLHBOX8 SUBFAMILY.  
-1- SIMILARITY: Contains 1 homeobox domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
-----  
CC EMBL U35632; AAA88820.1; -;  
DR EMBL S82178; AAB47101.1; -;  
DR EMBL S82168; AAB47101.1; JOINED.  
DR EMBL U30329; AAB47101.1; -;  
DR EMBL X99894; CAA68169.1; -;  
DR EMBL AF035250; AAB88463.1; -;  
DR EMBL AF035258; AAB88463.1; JOINED.  
DR EMBL AF049893; AAC05157.1; -;  
DR PIR G01926; G01926.  
DR HSSP P02833; 9ANT.  
DR TRANSFAC T04362; -;  
DR GeneW HGNC:6107; IPF1.  
DR MIM 600733; -;  
DR MIM 260370; -;  
DR MIM 125853; -;  
DR MIM 606392; -;  
DR GO GO:0003704; F:specific RNA polymerase II transcription fa. .; TAS.  
DR GO GO:0006091; P:energy pathways; TAS.  
DR GO GO:0007397; P:histogenesis and organogenesis; TAS.  
DR InterPro IPR001827; Antennapedia.  
DR InterPro IPR001356; Homeobox.  
DR InterPro IPR00047; HTH\_LambdaRepressor.  
DR Pfam PF00046; homeobox; 1.  
DR PRINTS PR00025; ANTENNAPEDIA.  
DR PRINTS PR00024; HOMEBOX.  
DR PRINTS PR00031; HTRREPRESSR.  
DR PRODOM PR000010; Homeobox; 1.  
DR SMART SM00388; HOX; 1.  
DR PROSITE PS00027; HOMEBOX\_1; 1.  
DR PROSITE PS00071; HOMEBOX\_2; 1.  
DR PROSITE PS00032; ANTENNAPEDIA.  
DR Transcription regulation; Activator; DNA-binding; Homeobox;  
KW Nuclear protein; Developmental protein; Disease mutation;  
KW Diabetes mellitus; Phosphorylation; Polymorphism.  
FT DOMAIN 13 73 TRANSACTIVATION DOMAIN (BY SIMILARITY).  
FT 42 51 POLY-PRO.  
FT SITE 118 123 ANTP-TYPE HEXAPEPTIDE.  
FT DNA\_BIND 146 205 HOMEBOX.  
FT DOMAIN 197 203 NUCLEAR LOCALIZATION SIGNAL (BY  
FT SIMILARITY).  
FT DOMAIN 216 219 POLY-PRO.  
FT 239 244 POLY-GLY.  
FT VARIANT 18 18 C->R (IN MODY4).  
FT 59 59 /FTid=VAR\_009309.  
FT 76 76 Q->L (IN MODY4).  
FT 76 76 D->N (IN MODY4; COULD BE A  
FT POLYMORPHISM).  
FT 197 197 /FTid=VAR\_009311.  
FT 197 197 R->H (IN MODY4).  
FT 243 243 /FTid=VAR\_009312.  
FT 243 243 P->P (IN MODY4).  
FT 243 243 /FTid=VAR\_009313.  
FT 56 56 A->S (IN REF. 6).  
FT 116 116 Q->H (IN REF. 6).  
FT 210 211 GG->SS (IN REF. 6).  
SQ SEQUENCE 283 AA; 30771 MW; 5048EC8D3289F72B CRC64;  
Query Match 10.9%; Score 69; DB 1; Length 283;  
Best Local Similarity 27.7%; Pred. No. 14;  
Matches 28; Conservative 7; Mismatches 38; Indels 28; Gaps 5;



QY 6 ANGEIVQDDDPVRRVTTTQPPRGSIIPROSPFNHGA-----PGGPGPQQOAGARLGA 59  
 DB 53 ALGLEQSGPPDISPYEVPFLAEDPAVAHLHHLPAQLGLAHPFGSPFNGTEPGGLEEP 108  
 QY 60 GAQSPFNDLNRLQVVMGFP-----QWHLGNHAYEP 88  
 DB 109 LEEP-----NR--VOLPFPWKKSTRAHAWKQWAGAYAAEP 143

RESULT 15  
 ID IPFL\_MESAU STANDARD; PRT; 283 AA.  
 AC P70118;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Insulin promoter factor 1 (IPF-1) (Homeobox domain protein PDX1).  
 GN IPF1 OR PDX1.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rudnick A., Ling T.Y., Odagiri H., Rutter W.J., German M.S.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Activates insulin and somatostatin gene transcription.  
 CC Key regulator of islet peptide hormone expression but also  
 CC responsible for the development of the pancreas, most probably by  
 CC determining maturation and differentiation of common pancreatic  
 CC precursor cells in the developing gut. Binds the DNA sequence 5'-  
 CC CC[CT]TATGGG-3' (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -1- TISSUE SPECIFICITY: Pancreas; Islet beta-cells (By similarity).  
 CC -1- DOMAIN: The Anp-type hexapeptide mediates heterodimerization of  
 CC IPF1 with PDX1 on a regulatory element of the somatostatin promoter  
 CC (By similarity).  
 CC -1- DOMAIN: The homeobox domain, which contains the nuclear localization  
 CC signal, not only mediates DNA-binding, but also acts as a protein-  
 CC protein interaction domain for TCF3 (E47), NEUROD1 and HMG-I(Y) (By  
 CC similarity).  
 CC -1- SIMILARITY: BELONGS TO THE ANTP HOMEBOX FAMILY.  
 CC IPF1/XLHBOX SUPERFAMILY.  
 CC -1- SIMILARITY: Contains 1 homeobox domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U73854; AAB18252.1; -.  
 DR HSSP; P02833; 9ANT.  
 DR TRASNSEC; T04267; -.  
 DR InterPro: IPR001827; Antennapedia.  
 DR InterPro: IPR001355; Homeobox.  
 DR InterPro: IPR000047; HTH\_1amdrepressr.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00025; ANTENNAPEDIA.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOK; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEBOX\_2; 1.  
 DR PROSITE; PS00032; ANTENNAPEDIA; FALSE NEG.  
 KM Transcription regulation; Activator; DNA-binding; Homeobox;  
 KM Nuclear protein.  
 FT DOMAIN 13 73 TRANSACTIVATION DOMAIN (BY SIMILARITY).  
 FT 42 48 POLY-PRO.

FT SITE 118 123 ANTP-TYPE HEXAPEPTIDE.  
 FT DNA BIND 146 205 HOMEBOX.  
 FT DOMAIN 197 203 NUCLEAR LOCALIZATION SIGNAL (BY  
 FT SIMILARITY).  
 FT POLY-PRO.  
 SQ SEQUENCE 283 AA; 30819 MW; 22714BBSARD54BF2 CRC64;  
 QY 6 ANGEIVQDDDPVRRVTTTQPPRGSIIPROSPFNHGA-----PGGPGPQQOAGARLGA 57  
 DB 53 ALGLEQSGPPDISPYEVPFLAEDPAVAHLHHLPAQLGLAHPFGSPFNGTEPGGLEEP 112  
 QY 58 GAQSPFNDLNRLQVVMGFP-----QWHLGNHAYEP 88  
 DB 113 SRQQLPFPWKKSTRAHAWKQWAGAYAAEP 143

Search completed: August 24, 2004, 09:18:24  
 Job time : 12 secs

Blank

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:04 ; Search time 40 Seconds

(without alignments)  
954,443 Million cell updates/sec

Title: US-10-001-885-125

Sequence: 1 MVRILANGELIVQDDPRVRT.....GVKGLLVGLVYVSHLSQR 121

Scoring table: BLOSUM62

Gap: 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 632   | 100.0       | 121    | 4     | Q96D05      |
| 2          | 568.5 | 90.0        | 120    | 11    | Q9D882      |
| 3          | 118.5 | 18.8        | 131    | 11    | Q9CZL2      |
| 4          | 110.5 | 17.5        | 132    | 4     | Q8N877      |
| 5          | 110   | 17.4        | 132    | 4     | Q8T859      |
| 6          | 99.5  | 15.7        | 223    | 16    | Q8ZHR5      |
| 7          | 83.5  | 13.2        | 219    | 4     | Q9NT75      |
| 8          | 83    | 13.1        | 174    | 10    | Q7XRN3      |
| 9          | 82    | 13.0        | 608    | 12    | Q9O5K9      |
| 10         | 82    | 13.0        | 475    | 16    | Q9A913      |
| 11         | 81    | 12.8        | 392    | 16    | Q8Z0Y1      |
| 12         | 80    | 12.7        | 717    | 11    | Q8C695      |
| 13         | 78.5  | 12.4        | 1318   | 16    | Q7UG24      |
| 14         | 78    | 12.3        | 102    | 12    | Q36158      |
| 15         | 78    | 12.3        | 102    | 12    | Q36158      |
| 16         | 78    | 12.3        | 102    | 12    | Q39945      |

|    |      |      |      |    |        |                     |
|----|------|------|------|----|--------|---------------------|
| 17 | 77.5 | 12.3 | 692  | 5  | Q19579 | Q19579 caenorhabd   |
| 18 | 76.5 | 12.1 | 214  | 12 | Q999V2 | Q999V2 hepatitis d  |
| 19 | 76.5 | 12.1 | 1201 | 16 | Q9H70  | Q9H70 pseudomonas   |
| 20 | 76   | 12.0 | 102  | 12 | Q36114 | Q36114 hepatitis d  |
| 21 | 76   | 12.0 | 140  | 11 | Q9DCB6 | Q9DCB6 mus musculus |
| 22 | 76   | 12.0 | 204  | 3  | Q13598 | Q13598 schistosach  |
| 23 | 76   | 12.0 | 252  | 4  | Q9NKS3 | Q9NKS3 mus musculus |
| 24 | 76   | 12.0 | 252  | 4  | Q81XT8 | Q81XT8 mus musculus |
| 25 | 76   | 12.0 | 395  | 11 | Q8BY82 | Q8BY82 mus musculus |
| 26 | 76   | 12.0 | 717  | 11 | Q9CXJ4 | Q9CXJ4 mus musculus |
| 27 | 76   | 12.0 | 717  | 11 | Q8C7W4 | Q8C7W4 mus musculus |
| 28 | 75   | 11.9 | 102  | 12 | Q36127 | Q36127 hepatitis d  |
| 29 | 75   | 11.9 | 102  | 12 | Q36153 | Q36153 hepatitis d  |
| 30 | 75   | 11.9 | 102  | 12 | Q36128 | Q36128 hepatitis d  |
| 31 | 75   | 11.9 | 701  | 4  | Q8N1X3 | Q8N1X3 mus musculus |
| 32 | 74.5 | 11.8 | 426  | 2  | Q87638 | Q87638 pseudomonas  |
| 33 | 74.5 | 11.8 | 1003 | 5  | Q9W326 | Q9W326 drosophila   |
| 34 | 74.5 | 11.8 | 1151 | 11 | Q9Z0F2 | Q9Z0F2 mus musculus |
| 35 | 74   | 11.7 | 102  | 12 | Q36121 | Q36121 hepatitis d  |
| 36 | 74   | 11.7 | 214  | 12 | Q9E925 | Q9E925 hepatitis d  |
| 37 | 74   | 11.7 | 256  | 2  | Q9F5M2 | Q9F5M2 rhodocycilus |
| 38 | 74   | 11.7 | 588  | 4  | Q9BWL8 | Q9BWL8 mus musculus |
| 39 | 74   | 11.7 | 668  | 16 | Q881Q1 | Q881Q1 pseudomonas  |
| 40 | 74   | 11.7 | 1260 | 4  | Q9UGY9 | Q9UGY9 mus musculus |
| 41 | 74   | 11.7 | 1794 | 12 | P87515 | P87515 barmah fore  |
| 42 | 73.5 | 11.6 | 81   | 12 | Q91Z80 | Q91Z80 hepatitis d  |
| 43 | 73.5 | 11.6 | 195  | 12 | Q9E924 | Q9E924 hepatitis d  |
| 44 | 73.5 | 11.6 | 254  | 5  | Q9N357 | Q9N357 caenorhabd   |
| 45 | 73.5 | 11.6 | 332  | 11 | Q9D7D1 | Q9D7D1 mus musculus |

#### ALIGNMENTS

RESULT 1  
ID Q96D05 PRELIMINARY; PRT: 121 AA.

AC Q96D05; 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Similar to RIKEN CDNA 2010107G23 gene.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC013587; AAH13587.1; -  
SQ SEQUENCE 121 AA; 13238 MW; 6326728CAB5375C4 CRC64;

Query Match 100.0%; Score 632; DB 4; Length 121;  
Best Local Similarity 100.0%; Pred. No. 1e-60;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |                                                            |     |
|----|-----|------------------------------------------------------------|-----|
| QY | 1   | MVRILANGELIVQDDPRVRTTTTPPPRSIPROSPFNRGHAGAPPGGPGRQOAGRLGAA | 60  |
| DB | 1   | MVRILANGELIVQDDPRVRTTTTPPPRSIPROSPFNRGHAGAPPGGPGRQOAGRLGAA | 60  |
| QY | 61  | QSPFNDLRQVLVNGGFPOMHIGNHAEVPTSIILLFILMIGVGLIVGVYVSHLSQ     | 120 |
| DB | 61  | QSPFNDLRQVLVNGGFPOMHIGNHAEVPTSIILLFILMIGVGLIVGVYVSHLSQ     | 120 |
| QY | 121 | R                                                          | 121 |
| DB | 121 | R                                                          | 121 |

RESULT 2  
Q9D882

ID Q9D882 PRELIMINARY; PRT; 120 AA.  
 AC Q9D882;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE 2010107G23Rik protein (RIKEN cDNA 2010107G23 gene) (Hypothetical protein).  
 GN 2010107G23Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Small intestine;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S., Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 RL Nature 409:685-690(2001).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;  
 RA Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK008338; BAB25613.1; -  
 DR EMBL; BC024943; AAH24943.1; -  
 DR EMBL; AK035688; BAC29152.1; -  
 DR MGD; MGI:1917144; 2010107G23Rik.  
 KW Hypothetical protein.  
 SQ SEQUENCE 120 AA; 13118 MW; 92C337D6CC6B4460 CRC64;

Query Match 90.0%; Score 568.5; DB 11; Length 120;  
 Best Local Similarity 92.6%; Pred. No. 7.8e-54;  
 Matches 112; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MVRILANEIVQDDPRVRRITTPPGSGIPROSGFNRGCHGAPGCGPPOOQAGARLGA 60  
 DB 1 MVRILANEIVQDDPRVRRITTPPGSGIPROSGFNRGCHGAPGCGPPOOQAGARLGA 59  
 QY 61 OSPFNDLNRQIVNNGFPQWHLGNHAEVPTVSIILFLIMLGVAGLLVGLVAVLSHQ 120  
 DB 60 QSPFSDLNRQIVNNGFPQWHLGNHAEVPTVSIILFLIMLGVAGLLVGLVAVLSHQ 119  
 QY 121 R 121  
 DB 120 R 120

RESULT 3  
 ID Q9CZL2 PRELIMINARY; PRT; 131 AA.  
 AC Q9CZL2;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE 2700063A17Rik protein.  
 GN 5730508B09Rik OR 2700063A17Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S., Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK012470; BAB28264.1; -  
 DR MGD; MGI:1917867; 5730508B09Rik.  
 SQ SEQUENCE 131 AA; 14511 MW; 2F4937AADFF169 CRC64;

Query Match 18.8%; Score 118.5; DB 11; Length 131;  
 Best Local Similarity 28.6%; Pred. No. 5.3e-05;  
 Matches 34; Conservative 19; Mismatches 49; Indels 17; Gaps 3;

QY 6 ANGEIVQDDPRVRRITTPPGSGIPROSGFNRGCHGAPGCGPPOOQAGARLGA 58  
 DB 15 AGCEDEDEGAPAGAVEPHERGASPR-----RRQDGEDEIDEPQNHSGSPIDD 66  
 QY 59 --AOSPFNDLNRQIVNNGFPQWHLGNHAEVPTVSIILFLIMLGVAGLLVGLVAVLSHQ 115  
 DB 67 YKMGTLFGELINKILNNGFTFRMYGERIVPEPVVTLFFMLMLWELGQALGVAVLCLV 125

RESULT 4  
 ID Q8N8U7 PRELIMINARY; PRT; 132 AA.  
 AC Q8N8U7;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein FLJ39370.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RP SEQUENCE FROM N.A.  
 RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K., Arita M., Mutsaers K., Yuuki H., Hara H., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto U., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,

RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,  
 RA Sugano S., Nagahara K., Masubo Y., Nagai K., Isogai T.,  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK096689; BAC04841.1; -  
 KM Hypothetical protein.  
 SQ SEQUENCE 132 AA; 14725 MW; D241F43200E41FDE CRC64;  
 Query Match 17.5%; Score 110.5; DB 4; Length 132;  
 Best Local Similarity 31.9%; Pred. No. 0.00039;  
 Matches 30; Conservative 17; Mismatches 34; Indels 13; Gaps 3;  
 QY 26 RGSIPRSG---FNRHGHAPPGSGPFGPQQAGARLGAASPFNDLRQLVNMGPQWHL 81  
 DB 42 RGSIPRSGEDVEDSDQHTGEPVGDYK-----KMGTL---FGLNKNLNMGFTRMYP 92  
 QY 82 GNHAEVPTSTLLFLMMGVRLVGLVYLV 115  
 DB 93 GERIVEPVIVIFFWMLMFLGLQALGLVAVLCIV 126  
 RESULT 5  
 Q8TRF9 PRELIMINARY; PRT; 132 AA.  
 AC Q8TRF9;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Strauberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC022534; A422534.1; -  
 KM Hypothetical protein.  
 SQ SEQUENCE 132 AA; 14787 MW; 284B5499A1E50P05 CRC64;  
 Query Match 17.4%; Score 110; DB 4; Length 132;  
 Best Local Similarity 29.3%; Pred. No. 0.00044;  
 Matches 27; Conservative 14; Mismatches 29; Indels 22; Gaps 1;  
 QY 46 PGRPOQAGARLGAASPFNDLRQLVNMGPQWHL 83  
 DB 35 PGRPOQAGARLGAASPFNDLRQLVNMGPQWHL 83  
 QY 84 HAVPVTSTLLFLMMGVRLVGLVYLV 115  
 DB 95 RIVERPVIVIFFWMLMFLGLQALGLVAVLCIV 126  
 RESULT 6  
 Q82HR5 PRELIMINARY; PRT; 223 AA.  
 AC Q82HR5;  
 DT 01-JUN-2003 (TREMblrel. 24, Created)  
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 NC NCB1\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis."  
 RL Nat. Biotechnol. 21:526-531 (2003).  
 DR EMBL; AP005035; BAC71155.1; -  
 KM Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 223 AA; 23128 MW; AF13F56F4DB3A87 CRC64;  
 Query Match 15.7%; Score 99.5; DB 16; Length 223;  
 Best Local Similarity 30.9%; Pred. No. 0.011;  
 Matches 29; Conservative 9; Mismatches 21; Indels 35; Gaps 4;  
 QY 23 QPP---RGSIPRSG---FNRHGHAP--PGSGPFGPQQAGARLGAASPFNDLRQL 71  
 DB 27 QPPYQAPRGIPOGIGYPOGQPGYGPGGPFGPQVASKGRRLGA----- 74  
 QY 72 VNMGPQWHLGNHAEVPTSTLLFLMMGVRLVGLVYLV 105  
 DB 75 -----RAIDGVAFFVIFLGLVAVLCIV 96  
 RESULT 7  
 Q9NT75 PRELIMINARY; PRT; 219 AA.  
 AC Q9NT75;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN DKFZP434N2030.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL137488; CAB70763.1; -  
 DR PIR; T46472; T46472.  
 KM Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 219 AA; 24989 MW; EB4D7AC00C28F97B CRC64;  
 Query Match 13.2%; Score 83.5; DB 4; Length 219;  
 Best Local Similarity 36.5%; Pred. No. 0.6;  
 Matches 23; Conservative 11; Mismatches 24; Indels 5; Gaps 2;  
 QY 41 APPGPGPQQAGARLGAASPFNDLRQLVNMGPQWHLGNHAEVPTSTLLFLMM 100  
 DB 9 APPRAGSVKTRRG-RGLNHLIPC---RPGISVGFSGEVMGSGSVRELOVVLFLVYR 63  
 QY 101 LGV 103  
 DB 64 EGI 66  
 RESULT 8  
 Q7XRN3 PRELIMINARY; PRT; 174 AA.  
 AC Q7XRN3;  
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 NC NCB1\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE OSJNB0024J22.19 protein.  
GN OSJNB0024J22.19.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriocaridaceae; Oryzaeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
RA Liu Y.L., Mu J., Xu Z., Chen L., Fan D.L., Wang Q.J., Zhang L.,  
RA Lu Y.Q., Yu S.L., Liu X.H., Li T.T., Zhang Y.Q., Lu Y., Li C., Li T.,  
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,  
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.,  
RA Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AL71596; CA02411.1; -  
SQ SEQUENCE 174 AA; 17749 MW; 6BD95E3BD5117DC7 CRC64;

Query Match 13.1%; Score 83; DB 10; Length 174;  
Best Local Similarity 32.1%; Pred. No. 0.52; Mismatches 23; Indels 24; Gaps 3;  
Matches 27; Conservative 10; Mismatches 23; Indels 24; Gaps 3;  
DB 14 GCGAPTCAG-----SGAGAGAPGAVNDNN-----CGNSAPSSGGLPSGY 90  
QY 38 GHGAPGPGPRQOAGARLGAASPFNDLRQLVNGFPMWLGNAVEPVT----- 90  
DB 14 GCGAPTCAG-----SGAGAGAPGAVNDNN-----CGNSAPSSGGLPSGY 90  
QY 91 SILLFLMLGVKGLIVGLYL 114  
DB 57 NILLFLMLFLPPLALSLYYV 80

RESULT 9  
ID 036094 PRELIMINARY; PRT; 102 AA.  
AC 036094;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hepatitis delta antigen (Fragment).  
OS Hepatitis delta virus (HDV).  
OC Viruses; Deltavirus.  
OX NCBI\_TaxID=12475;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Turkish isolate 07;  
RX MEDLINE=97378276; PubMed=9734957;  
RA Shail A.O., Haddiyannis S., Hootnagle J.H., Di Bisceglie A.M.,  
RA Gerin J.L., Casey J.L.,  
RT "Geographic distribution and genetic variability of hepatitis delta  
RT virus genotype 1.";  
RL Virology 234:160-167(1997).  
DR EMBL: AF008349; AAB6964.1; -  
DR GO: 0042025; C:host cell nucleus; IEA.  
DR GO: 0003723; F:RNA binding; IEA.  
DR InterPro: IPR002506; HDV\_ag;  
DR Pfam: PF01517; HDV\_ag; 1.  
DR Prodom: PD002887; HDV\_ag; 1.  
FT NON TER 1  
SQ SEQUENCE 102 AA; 11085 MW; 3136BB4EA439F792 CRC64;

Query Match 13.0%; Score 82; DB 12; Length 102;  
Best Local Similarity 32.9%; Pred. No. 0.36; Mismatches 27; Indels 14; Gaps 5;  
Matches 25; Conservative 10; Mismatches 27; Indels 14; Gaps 5;

DB 21 LTERERRERRTAGPPTGVNPSB---GERRGAPGGFVPSKQ-----GVPSFPHRLG 71  
QY 69 RQL---VAMGFPQWHL 81  
DB 72 EGLDVKSGSGFP-WDI 86

RESULT 10  
ID 0905K9 PRELIMINARY; PRT; 608 AA.  
AC 0905K9;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE NTR.  
OS Herpesvirus papio.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gamaherpesvirinae; Lymphocryptovirine.  
OX NCBI\_TaxID=10394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Baboon lymphocryptovirus BA65;  
RA Zeng J.-C., Ryan J., Ling F.D., Loeb D.D., Pagano J.S., Hayward S.D.,  
RA Hayward G.S.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF200364; AAF23950.1; -  
SQ SEQUENCE 608 AA; 60458 MW; 4BF82ACB0A029627 CRC64;

Query Match 13.0%; Score 82; DB 12; Length 608;  
Best Local Similarity 47.4%; Pred. No. 2.8; Mismatches 17; Indels 0; Gaps 0;  
Matches 18; Conservative 3; Mismatches 17; Indels 0; Gaps 0;  
DB 65 PPGAGQRPSPGPTGCHPAPGAPRSPRTRRRGSAG 102  
QY 24 PPGSIPRSGPFRNGHAPGPGPRQOAGARLGAAG 61  
DB 65 PPGAGQRPSPGPTGCHPAPGAPRSPRTRRRGSAG 102

RESULT 11  
ID 09A913 PRELIMINARY; PRT; 475 AA.  
AC 09A913;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Sensor histidine kinase.  
GN CCL181.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
OC Caulobacteraceae; Caulobacter.  
OX NCBI\_TaxID=155892;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R., Ohta N., Maddock J.R.,  
RA Potocka I., Nelson W.C., Newton R.J., Durkin A.S., Gwinn M.L., Harte D.H.,  
RA Kolony J.F., Smit J., Craven M.B., Khoult H., Shetty J., Berry K.,  
RA Utecherback T., Tran K., Wolf A., Vamathavan J., Ernoelava M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
CC -1- SIMILARITY TO PROKARYOTE SENSOR TRANSDUCTION PROTEINS.  
DR EMBL: AB005795; AAK23165.1; -  
DR PIR: A87396; A87396.  
DR HSSP: P02933; 1BXD.  
DR TIGR: CCL181; -  
DR GO: 0016020; C:membrane; IEA.  
DR GO: 0005524; F:ATP binding; IEA.  
DR GO: 0016301; F:kinase activity; IEA. TT

DR GO; GO:0007600; P:sensory perception; IEA.  
DR GO; GO:0007165; P:signal transduction; IEA.  
DR InterPro; IPR003594; ATPbind ATPase.  
DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
DR InterPro; IPR003660; HAMF.  
DR InterPro; IPR005467; His\_kinase.  
DR InterPro; IPR003661; His\_kin\_N.  
DR Pfam; PF00672; HAMF\_1.  
DR Pfam; PF02518; HAMFase\_C\_1.  
DR Pfam; PF00512; HisKA\_1.  
DR PRINTS; PR00344; BCTRLSENSOR.  
DR SMART; SM00304; HAMF\_1.  
DR SMART; SM00387; HAMFase\_C\_1.  
DR SMART; SM00388; HisKA\_1.  
DR PROSITE; PS50885; HAMF\_1.  
DR PROSITE; PS50109; HIS\_KIN\_1.  
DR Kinase; Phosphorylation; Sensory transduction; Transferase;  
KW Complete proteome.  
SQ SEQUENCE 475 AA; 51395 MW; 5178BCAF86706349 CRC64;  
  
Query Match 12.8%; Score 81; DB 16; Length 475;  
Best Local Similarity 25.0%; Pred. No. 2.7;  
Matches 36; Conservative 9; Mismatches 39; Indels 40; Gaps 7;  
  
QY 9 EIVQDDPRV---RTTQPRGSI-----RQSFNKGHPGPGPCROQA--53  
DB 104 ELQJDNQPRFVRRTPTPRPDAPAGANGRPAGRPAREPAREG-GAPCGDNARSQARD 162  
QY 54 -----GARLGAAGSPFNDLNROLV-----NMGPQHLGNHVEPTSTILLFLMLTG 102  
DB 163 EPLIFGDFKVGKVRQSD-----GRWLVEPKPTLRDSWQ-----QRILLILLSVIA 209  
QY 103 VRGL 106  
DB 210 VTPL 213  
  
RESULT 12  
082YV1  
ID 082YV1 PRELIMINARY; PRT; 392 AA.  
AC 082YV1;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative phosphatidate cytidyllyltransferase.  
GN CDSA OR SAV2623.  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33903;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=21477403; PubMed=11572948;  
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
RT "Genome sequence of an industrial microorganism Streptomyces  
RT avermitilis: deducing the ability of producing secondary  
RT metabolites";  
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
RL [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=22608306; PubMed=12692562;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
RA Sakaki Y., Hattori M., Omura S.;  
RT "Complete genome sequence and comparative analysis of the industrial  
RT microorganism Streptomyces avermitilis";  
RT Nat. Biotechnol. 21:526-531(2003).  
RL EMBL; AP005031; BAC70334.1;  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004605; F:phosphatidate cytidyllyltransferase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.  
DR InterPro; IPR000374; PC\_trans.  
DR Pfam; PF01148; CRP\_transf\_1.  
DR PROSITE; PS01315; CDS\_1.  
KW Nucleotidyltransferase; Transferase; Complete proteome.  
SQ SEQUENCE 392 AA; 40876 MW; C6F6A4254E7C9AC6 CRC64;  
  
Query Match 12.7%; Score 80; DB 16; Length 392;  
Best Local Similarity 27.2%; Pred. No. 2.8;  
Matches 37; Conservative 18; Mismatches 33; Indels 48; Gaps 8;  
  
QY 7 NGEIVQDDPRV---RTTQPRGSI-----RQSFNKGHPGPGPCROQA--53  
DB 67 SGPLFRDDTPQTTSGSQPTVYTGSPYGS-----QTPHGAPQKPEMPAPQAPAP 119  
QY 49 RQOQAGARLGAAGSPFNDLNROLVNMGPQHLGNHVEPTSTILLFL--TMLGVRL 106  
DB 120 OKXAGRDGLGAIG-----VGVG-----VIVASLFFVKAVFVGLAV 159  
QY 107 -LLNGLVYVSHLSOR 121  
DB 160 AVVGLWELTSRLER 175  
  
RESULT 13  
08C695  
ID 08C695 PRELIMINARY; PRT; 717 AA.  
AC 08C695;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ATP-binding cassette.  
GN 4833412N02RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RX MEDLINE=22354683; PubMed=1246851;  
RA The FANTOM Consortium;  
RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RT Nature 420:563-573(2002).  
RL EMBL; AK076315; BAC36297.1;  
DR MGD; MGI:1921860; 4833412N02RIK.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; P:ATP binding; IEA.  
DR GO; GO:0004009; P:ATP-binding cassette (ABC) transporter acti. .; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR001140; ABC\_TM\_transp.  
DR InterPro; IPR003439; ABC\_transporter.  
DR Pfam; PF00664; ABC\_membrane\_1.  
DR Pfam; PF00005; ABC\_tran\_1.  
DR Prodom; PD000006; ABC\_transporter\_1.  
DR SMART; SM00382; AAA\_1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
SQ SEQUENCE 717 AA; 77983 MW; 72CC6DE0C967755 CRC64;  
  
Query Match 12.7%; Score 80; DB 11; Length 717;  
Best Local Similarity 37.1%; Pred. No. 5.6;  
Matches 26; Conservative 13; Mismatches 25; Indels 6; Gaps 3;  
  
QY 54 GARLGAAGSPFNDLNROLVNM--GPPQHLGNHVEPTSTILLFLMLGVRL 111  
DB 139 GAALVNVQGPL--LLGQVEMVAKYTRDHMSFVSE--SRKLSVQLLLYGVQGLTFEGY 194





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:06 ; Search time 19 Seconds

(without alignments)  
328.776 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632  
Sequence: 1 MRRILANGELIVQDDPRVPT.....GVAGLLLVGLVIVHSLSQR 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptcdat2/2/1aa/5A COMB.pep.\*
- 2: /cgn2\_6/ptcdat2/2/1aa/5B COMB.pep.\*
- 3: /cgn2\_6/ptcdat2/2/1aa/6A COMB.pep.\*
- 4: /cgn2\_6/ptcdat2/2/1aa/6B COMB.pep.\*
- 5: /cgn2\_6/ptcdat2/2/1aa/PTCUS COMB.pep.\*
- 6: /cgn2\_6/ptcdat2/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Length | DB ID | Description                            |
|------------|-------|--------------|-------|----------------------------------------|
| 1          | 551   | 87.2         | 104   | US-09-621-976-4724 Sequence 4724, Ap   |
| 2          | 387   | 61.2         | 111   | US-09-149-476-673 Sequence 673, App    |
| 3          | 245   | 38.8         | 113   | US-09-149-476-673 Sequence 367, App    |
| 4          | 240   | 38.0         | 113   | US-09-621-976-6924 Sequence 6924, Ap   |
| 5          | 88    | 13.9         | 460   | US-09-489-039A-8966 Sequence 8966, Ap  |
| 6          | 79.5  | 12.6         | 763   | US-09-252-991A-30146 Sequence 30146, A |
| 7          | 78.5  | 12.4         | 605   | US-08-889-402-2 Sequence 2, Appl       |
| 8          | 76    | 12.0         | 217   | US-09-252-991A-23015 Sequence 23015, A |
| 9          | 75    | 11.9         | 200   | US-09-252-991A-22303 Sequence 22303, A |
| 10         | 75    | 11.9         | 655   | US-09-252-991A-25314 Sequence 25314, A |
| 11         | 74.5  | 11.8         | 215   | US-09-252-991A-31559 Sequence 31559, A |
| 12         | 74    | 11.7         | 263   | US-09-252-991A-28276 Sequence 28276, A |
| 13         | 73.5  | 11.6         | 638   | US-09-252-991A-24325 Sequence 24325, A |
| 14         | 73    | 11.6         | 270   | US-09-252-991A-22026 Sequence 22026, A |
| 15         | 73    | 11.6         | 351   | US-09-252-991A-32396 Sequence 32396, A |
| 16         | 72.5  | 11.5         | 482   | US-09-252-991A-31705 Sequence 31705, A |
| 17         | 71.5  | 11.3         | 246   | US-09-252-991A-28637 Sequence 28637, A |
| 18         | 71.5  | 11.3         | 415   | US-09-252-991A-23751 Sequence 23751, A |
| 19         | 71.5  | 11.3         | 427   | US-09-252-991A-31904 Sequence 31904, A |
| 20         | 71.5  | 11.3         | 734   | US-09-252-991A-33036 Sequence 33036, A |
| 21         | 71.5  | 11.3         | 1053  | US-09-252-991A-26140 Sequence 26140, A |
| 22         | 71    | 11.2         | 199   | US-08-122-458D-18 Sequence 18, Appl    |
| 23         | 71    | 11.2         | 298   | US-08-122-458D-11 Sequence 11, Appl    |
| 24         | 71    | 11.2         | 480   | US-09-252-991A-23424 Sequence 23424, A |
| 25         | 71    | 11.2         | 645   | US-09-252-991A-22095 Sequence 22095, A |
| 26         | 71    | 11.2         | 988   | US-09-252-991A-29659 Sequence 29659, A |
| 27         | 70.5  | 11.2         | 149   | US-09-252-991A-22850 Sequence 22850, A |

|    |      |      |     |                                        |
|----|------|------|-----|----------------------------------------|
| 28 | 70.5 | 11.2 | 295 | US-09-252-991A-22439 Sequence 22439, A |
| 29 | 70.5 | 11.2 | 390 | US-09-252-991A-22732 Sequence 22732, A |
| 30 | 70.5 | 11.2 | 531 | US-09-252-991A-24404 Sequence 24404, A |
| 31 | 70.5 | 11.2 | 863 | US-09-252-991A-21831 Sequence 21831, A |
| 32 | 70   | 11.1 | 253 | US-09-252-991A-27487 Sequence 27487, A |
| 33 | 70   | 11.1 | 380 | US-09-252-991A-22319 Sequence 22319, A |
| 34 | 69.5 | 11.0 | 324 | US-09-252-991A-24165 Sequence 24165, A |
| 35 | 69.5 | 11.0 | 329 | US-09-252-991A-18860 Sequence 18860, A |
| 36 | 69.5 | 11.0 | 410 | US-09-252-991A-25451 Sequence 25451, A |
| 37 | 69.5 | 11.0 | 526 | US-09-252-991A-24404 Sequence 24404, A |
| 38 | 69.5 | 11.0 | 579 | US-09-252-991A-30334 Sequence 30334, A |
| 39 | 69   | 10.9 | 243 | US-09-252-991A-25431 Sequence 25431, A |
| 40 | 69   | 10.9 | 478 | US-09-252-991A-24169 Sequence 24169, A |
| 41 | 69   | 10.9 | 871 | US-09-252-991A-19431 Sequence 19431, A |
| 42 | 68.5 | 10.8 | 225 | US-09-252-991A-16928 Sequence 16928, A |
| 43 | 68.5 | 10.8 | 305 | US-09-252-991A-27334 Sequence 27334, A |
| 44 | 68.5 | 10.8 | 406 | US-09-489-039A-12736 Sequence 12736, A |
| 45 | 68.5 | 10.8 | 555 | US-09-252-991A-31521 Sequence 31521, A |

## ALIGNMENTS

RESULT 1  
US-09-621-976-4724 Application US/09621976  
Sequence 4724, Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Giordano, J.Y.  
FILE OF INVENTION: ESTs and Encoded Human Proteins:  
TITLE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 4724  
LENGTH: 104  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-621-976-4724

Query Match 87.2%; Score 551; DB 4; Length 104;  
Best Local Similarity 100.0%; Pred. No. 7.4e-57;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |    |                              |                 |              |    |
|----|----|------------------------------|-----------------|--------------|----|
| QY | 1  | MRRILANGELIVQDDPRVPTTQPPRGSI | PROSFNRRHGAPGGP | PROQOAGARIGA | 60 |
| DB | 1  | MRRILANGELIVQDDPRVPTTQPPRGSI | PROSFNRRHGAPGGP | PROQOAGARIGA | 60 |
| QY | 61 | QSPFNDLNQLVNMGFPQWHLGNHVEPVT | SILLFLMLGVR     | 104          |    |
| DB | 61 | QSPFNDLNQLVNMGFPQWHLGNHVEPVT | SILLFLMLGVR     | 104          |    |

RESULT 2  
US-09-149-476-673 Application US/09149476  
Sequence 673, Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
FILE OF INVENTION: 186 Human Secreted proteins  
TITLE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621

|    |                                 |                                         |
|----|---------------------------------|-----------------------------------------|
| 1  | EARLIER FILING DATE: 1997-03-07 | EARLIER APPLICATION NUMBER: 60/040, 626 |
| 2  | EARLIER FILING DATE: 1997-03-07 | EARLIER APPLICATION NUMBER: 60/040, 334 |
| 3  | EARLIER FILING DATE: 1997-03-07 | EARLIER APPLICATION NUMBER: 60/040, 336 |
| 4  | EARLIER FILING DATE: 1997-03-07 | EARLIER APPLICATION NUMBER: 60/040, 163 |
| 5  | EARLIER FILING DATE: 1997-03-07 | EARLIER APPLICATION NUMBER: 60/047, 600 |
| 6  | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 615 |
| 7  | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 597 |
| 8  | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 502 |
| 9  | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 633 |
| 10 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 583 |
| 11 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 617 |
| 12 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 618 |
| 13 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 503 |
| 14 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 592 |
| 15 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 581 |
| 16 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 584 |
| 17 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 500 |
| 18 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 587 |
| 19 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 492 |
| 20 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 598 |
| 21 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 613 |
| 22 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 582 |
| 23 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 596 |
| 24 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 612 |
| 25 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 632 |
| 26 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047, 601 |
| 27 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/043, 580 |
| 28 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 566 |
| 29 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 314 |
| 30 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 569 |
| 31 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 311 |
| 32 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 671 |
| 33 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 674 |
| 34 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 669 |
| 35 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 312 |
| 36 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043, 313 |
| 37 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/047, 614 |

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043, 578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047, 501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043, 670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056, 632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048, 964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057, 650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056, 884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057, 669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049, 610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061, 060  
EARLIER FILING DATE: 1997-10-02

Query Match 61.2%; Score 387; DB 4; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1, 1e-37;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTNGETVODDDPRVTRTTTPPGSIPROSFNKGAGAPGPGPQQOAGALGAA 60  
DB 39 MVRILNGETVODDDPRVTRTTTPPGSIPROSFNKGAGAPGPGPQQOAGALGAA 98  
QY 61 QSPFNDRQLV 73  
DB 99 QSPFNDRQLV 111

RESULT 3  
US-09-149-476-367  
Sequence 367, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Roegen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149, 476  
CURRENT FILING DATE: 1998-03-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040, 162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040, 333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038, 621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040, 626  
EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040, 334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040, 336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040, 163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047, 600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043, 580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 315

EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 38.8%; Score 245; DB 4; Length 48;  
Best Local Similarity 100.0%; Pred. No. 1,3e-21;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 74 MGPFQWHLGNHNAVEPVTSILLFILMPLGVGGLLVGVIVSHLSOR 121  
Db 1 MGPFQWHLGNHNAVEPVTSILLFILMPLGVGGLLVGVIVSHLSOR 48

RESULT 4  
US-09-621-976-6924  
; Sequence 6924, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; NUMBER OF SEQ ID NOS: 2000-07-21  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6924  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 97  
; OTHER INFORMATION: Xaa = Pro,Gln  
US-09-621-976-6924

Query Match 38.0%; Score 240; DB 4; Length 113;  
Best Local Similarity 83.9%; Pred. No. 1.6e-20;  
Matches 47; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

QY 1 MRLANGELVODDDPRVTTTQPPRGSIIPROSGFNGHGAAPG--GPGPQQOAG 54  
Db 43 MRLANGELVODDDPRVTTTQPPRGSIIPROSGFNGHGAAPGVLAAPASSRQVXG 98

RESULT 5  
US-09-489-039A-8966  
; Sequence 8966, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Berton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709,2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8966  
; LENGTH: 460  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8966

Query Match 13.9%; Score 88; DB 4; Length 460;  
Best Local Similarity 31.4%; Pred. No. 0.052;  
Matches 22; Conservative 17; Mismatches 25; Indels 6; Gaps 3;  
QY 56 RLGAOSPPNDLNQOLVM-GFQPMHGNVAVPTVITLLFLMLGVAGLLVGLVYL 114  
Db 213 RLAVASPPGDIINMIAITLIGIPMHVGTIPKMTISVILLIGLITLGDALIV-LVFI 271  
QY 115 ----VSHLSQ 120  
Db 272 QPVLINMSE 281

RESULT 6  
US-09-252-991A-30146  
; Sequence 30146, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30146  
; LENGTH: 763  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30146

Query Match 12.6%; Score 79.5; DB 4; Length 763;  
Best Local Similarity 39.0%; Pred. No. 0.99;  
Matches 23; Conservative 5; Mismatches 22; Indels 9; Gaps 3;  
QY 12 QDDDPVRYTTQPPRGSIIPROSGFNGHGA-----PFGG--PGRQOQAGARLGAASPPNDLNRLV 63  
Db 443 QGGDGRPRTPDERAGATPGQAL--RAQGAAGATLPGGGRAGRRRPAAGRBAAGP 499

RESULT 7  
US-08-889-402-2  
; Sequence 2, Application US/0889402

Patent No. 5811288  
; GENERAL INFORMATION:  
; TITLE OF INVENTION: NOVEL ISOFORM GENE FOR FOCAL ADHESION  
; TITLE OF INVENTION: PROTEIN PAXILIN  
; NUMBER OF SEQUENCES: 7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/889,402  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 178334/1996  
; FILING DATE: 08-JUL-1996  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 605 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; CELL TYPE: Monoblastic cell and placenta  
; CELL LINE: U937  
US-08-889-402-2

Query Match 12.4%; Score 78.5; DB 2; Length 605;  
Best Local Similarity 34.9%; Pred. No. 0.96;  
Matches 22; Conservative 5; Mismatches 29; Indels 7; Gaps 2;  
QY 16 PVRITTPPPRGSIIPROSGF---NRGHGAPGPGPGRQOQAGARLGAASPPNDLNRLV 72  
Db 312 PSLRSPDGP-----PCQPMAGAKTGSSPPGPPKPSQDLSMGLSDLNKLVAV 367  
QY 73 NMG 75  
Db 368 AKG 370

RESULT 8  
US-09-252-991A-23015  
; Sequence 23015, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23015  
; LENGTH: 217  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23015

Query Match 12.0%; Score 76; DB 4; Length 217;  
Best Local Similarity 32.9%; Pred. No. 0.49;  
Matches 23; Conservative 3; Mismatches 18; Indels 26; Gaps 4;  
QY 26 RGSIPROSGFNGHGAAPG-----GP-----GPRQOQAGARLGAASPPNDLNRLV 72  
Db 91 RAGVPAARS-RRRGHGRPRGCVADRAGTGAAGTAPAPRHLHDAGASLPGSDPP----- 140  
QY 73 NMGPPQMLG 82

Db 141 ---GPAWHRG 147

## RESULT 9

US-09-252-991A-22303  
; Sequence 22303, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22303  
; LENGTH: 200  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22303

Query Match 11.9%; Score 75; DB 4; Length 200;  
Best Local Similarity 33.3%; Pred. No. 0.58; 29; Indels 10; Gaps 2;

Matches 22; Conservative 5; Mismatches 29; Indels 10; Gaps 2;

15 DPRVTTTPPGSGIPRQSFNNGHAPGPGPGPQOAGALGAQSPFNDINROLVNM 74

40 DEHVRSGAHPRKRCRATV---HRRPPGEPDLPQVGPRAAARP-----RLAGL 89

75 GPPQWH 80

90 GVQGRH 95

## RESULT 10

US-09-252-991A-25314  
; Sequence 25314, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25314  
; LENGTH: 655  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25314

Query Match 11.9%; Score 75; DB 4; Length 655;  
Best Local Similarity 33.3%; Pred. No. 2.7; 30; Indels 20; Gaps 6;

Matches 29; Conservative 8; Mismatches 30; Indels 20; Gaps 6;

3 RILANGELIVQ--DDPFR---VTTTTPPGSGIPRQSFNNGHAPGPGPGPQOAGAL 57

516 RIVEGGEVMEVGDPRHARHPKVRQPAR-----RPPGRPPGAPAGSHR--RRP 563

58 GAQSPFNDINROLVNMGPQWHLGNH 84

564 GAARSPVH--RRQPTGAG-PAGHNGSH 587

## RESULT 11

US-09-252-991A-31559  
; Sequence 31559, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31559  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31559

Query Match 11.8%; Score 74.5; DB 4; Length 215;  
Best Local Similarity 40.9%; Pred. No. 0.73; 20; Indels 1; Gaps 1;

Matches 18; Conservative 5; Mismatches 20; Indels 1; Gaps 1;

23 OPPRSIPRQSFNNGHAPG-GPPRQOAGALGAQSPFN 65

98 RPPRALGRSARGRARRRPPVLPGRRAQLGGRPGAGQNNFN 141

## RESULT 12

US-09-252-991A-28276  
; Sequence 28276, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28276  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28276

Query Match 11.7%; Score 74; DB 4; Length 263;  
Best Local Similarity 41.7%; Pred. No. 1.1; 7; Indels 22; Gaps 6;

Matches 25; Conservative 6; Mismatches 7; Indels 22; Gaps 6;

15 DPRVTTTTPPGSGIPRQSFNNGHAPGPGP-----RQOQAGAR-LGAQSP 63

110 DPRLR---QPVDAQDPAGDPR-----RGHGA---GPGPMLRPPRRRTAQAQAPAEHP 158

## RESULT 13

US-09-252-991A-24325  
; Sequence 24325, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24325  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24325

; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24325  
; LENGTH: 638  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24325

Query Match 11.6%; Score 73.5; DB 4; Length 638;  
Best Local Similarity 32.1%; Pred. No. 3.9;  
Matches 26; Conservative 11; Mismatches 33; Indels 11; Gaps 4;

QY 17 RVRTTTPRGSIROSFFNRGHGAPPGGPPRQQA-----QAGARLGAAG--SPFND 66  
DB 281 RRRRLAQPFRPPARRA-RQRAYPAGTDPQPGGPPRTGPPGAGAPGAVRHCRPYAN 339

QY 67 LAROLVNMGFPQWHLGNHAYE 87  
DB 340 PARRLVALGRHRCVAVQVAE 360

RESULT 14  
US-09-252-991A-22026  
; Sequence 22026, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22026  
; LENGTH: 270  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22026

Query Match 11.6%; Score 73; DB 4; Length 270;  
Best Local Similarity 38.9%; Pred. No. 1.5;  
Matches 28; Conservative 3; Mismatches 23; Indels 18; Gaps 5;

QY 17 RVRTTTPRGSIROSFFNRGHGAPPGGPPRQQA-----RLGAQSPFN 65  
DB 27 RVHVSPPPPRRGSLRRRS---RRPAHPGQSPRQHPGSDRLRGADRDAGGQCCPRV 82

QY 66 DLNRQLY-NMGF 76  
DB 83 -LRQLEGNPGEF 92

RESULT 15  
US-09-252-991A-32396  
; Sequence 32396, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 32396  
; LENGTH: 351  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32396

Query Match 11.6%; Score 73; DB 4; Length 351;  
Best Local Similarity 30.7%; Pred. No. 2.1;  
Matches 27; Conservative 3; Mismatches 40; Indels 18; Gaps 3;

QY 19 RRTTTPRGSIROSFFNRGHGAPPGG-----RQQAARLGAAGSP 63  
DB 47 RLPVPAKGRPPQ---RRSGAPRAGAPGSPRVGDHPPLRRPATDPAGDRPP 103

QY 64 FNDLNRQLYNMGFPQWHLGNHAYEPTS 91  
DB 104 AHPVGPALRRRLRLHLGGRTHPAMS 131

Search completed: August 24, 2004, 09:19:53  
Job time : 20 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:19:19 ; Search time 124 Seconds  
(without alignments)  
306.652 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632

Sequence: 1 MVRILANGELIVQDDPRVRR.....GVRGLLVGLVIVSHLSQR 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 31425058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
  - 2: /cgn2\_6/ptodata/1/pubppaa/PCF\_NEW\_PUB.pep:\*
  - 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*
  - 4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*
  - 6: /cgn2\_6/ptodata/1/pubppaa/PCFUS\_PUBCOMB.pep:\*
  - 7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
  - 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
  - 9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
  - 10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
  - 11: /cgn2\_6/ptodata/1/pubppaa/US09\_PUB.pep:\*
  - 12: /cgn2\_6/ptodata/1/pubppaa/US09\_PUB.pep:\*
  - 13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
  - 14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
  - 15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
  - 16: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*
  - 17: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*
  - 18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

| Result No. | Score | Query Match | Length DB | ID                   | Description          |
|------------|-------|-------------|-----------|----------------------|----------------------|
| 1          | 632   | 100.0       | 121       | US-10-001-885-125    | Sequence 125, App    |
| 2          | 632   | 100.0       | 121       | US-10-076-747-132    | Sequence 122, App    |
| 3          | 387   | 61.2        | 111       | US-09-809-391-673    | Sequence 673, App    |
| 4          | 387   | 61.2        | 111       | US-09-882-171-673    | Sequence 673, App    |
| 5          | 387   | 61.2        | 111       | US-10-164-861-673    | Sequence 673, App    |
| 6          | 245   | 38.8        | 48        | US-09-809-391-367    | Sequence 367, App    |
| 7          | 245   | 38.8        | 48        | US-09-882-171-367    | Sequence 367, App    |
| 8          | 245   | 38.8        | 48        | US-10-164-861-367    | Sequence 367, App    |
| 9          | 110.5 | 17.5        | 132       | US-10-108-260A-3882  | Sequence 3882, App   |
| 10         | 85.5  | 15.7        | 223       | US-10-156-761-19880  | Sequence 19880, App  |
| 11         | 85.5  | 15.7        | 112       | US-10-424-599-174474 | Sequence 174474, App |
| 12         | 82    | 13.0        | 123       | US-10-437-963-198292 | Sequence 198292, App |
| 13         | 81    | 12.8        | 256       | US-10-425-114-60982  | Sequence 60982, App  |
| 14         | 80    | 12.7        | 92        | US-10-106-698-8371   | Sequence 8371, App   |
| 15         | 80    | 12.7        | 392       | US-10-156-761-10160  | Sequence 10160, App  |

|    |      |      |       |                      |                      |
|----|------|------|-------|----------------------|----------------------|
| 16 | 80   | 12.7 | 542   | US-10-437-963-145131 | Sequence 145131, App |
| 17 | 79.5 | 12.6 | 286   | US-10-437-963-187696 | Sequence 187696, App |
| 18 | 76.5 | 12.1 | 225   | US-10-437-963-177234 | Sequence 177234, App |
| 19 | 76.5 | 12.1 | 2246  | US-10-437-963-164016 | Sequence 164016, App |
| 20 | 76   | 12.0 | 331   | US-10-425-114-64460  | Sequence 64460, App  |
| 21 | 75.5 | 11.9 | 622   | US-10-240-145-133    | Sequence 133, App    |
| 22 | 75   | 11.9 | 701   | US-10-108-260A-2701  | Sequence 2701, App   |
| 23 | 73.5 | 11.5 | 236   | US-10-437-963-149016 | Sequence 149016, App |
| 24 | 72.5 | 11.5 | 137   | US-10-437-963-193889 | Sequence 193889, App |
| 25 | 72.5 | 11.5 | 440   | US-10-369-493-18063  | Sequence 18063, App  |
| 26 | 72   | 11.4 | 259   | US-10-425-114-70983  | Sequence 70983, App  |
| 27 | 72   | 11.4 | 1126  | US-10-437-963-183078 | Sequence 183078, App |
| 28 | 71.5 | 11.3 | 195   | US-10-029-366-32263  | Sequence 32263, App  |
| 29 | 71.5 | 11.3 | 279   | US-10-425-114-40428  | Sequence 40428, App  |
| 30 | 71.5 | 11.3 | 378   | US-10-437-963-150350 | Sequence 150350, App |
| 31 | 71.5 | 11.3 | 1336  | US-10-437-963-115864 | Sequence 115864, App |
| 32 | 71.5 | 11.3 | 19662 | US-10-084-846A-6     | Sequence 6, App      |
| 33 | 71   | 11.2 | 166   | US-10-156-761-10938  | Sequence 10938, App  |
| 34 | 71   | 11.2 | 642   | US-10-263-929-105    | Sequence 105, App    |
| 35 | 71   | 11.2 | 793   | US-10-029-366-32918  | Sequence 32918, App  |
| 36 | 70.5 | 11.2 | 140   | US-10-424-599-231854 | Sequence 231854, App |
| 37 | 70.5 | 11.2 | 248   | US-10-425-114-42547  | Sequence 42547, App  |
| 38 | 70.5 | 11.2 | 915   | US-10-437-963-156444 | Sequence 156444, App |
| 39 | 70.5 | 11.2 | 974   | US-10-437-963-168292 | Sequence 168292, App |
| 40 | 70.5 | 11.2 | 1734  | US-10-437-963-153570 | Sequence 153570, App |
| 41 | 70   | 11.1 | 123   | US-10-424-599-145465 | Sequence 145465, App |
| 42 | 70   | 11.1 | 214   | US-10-205-194-164    | Sequence 164, App    |
| 43 | 70   | 11.1 | 417   | US-10-437-963-122270 | Sequence 122270, App |
| 44 | 70   | 11.1 | 16    | US-10-369-493-3240   | Sequence 3240, App   |
| 45 | 70   | 11.1 | 15    |                      |                      |

### ALIGNMENTS

RESULT 1  
US-10-001-885-125  
Sequence 125, Application US/10001885  
Publication No. US20040058319A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Recipon, Hervé  
APPLICANT: Cafferey, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Proteins  
FILE REFERENCE: DEX-0279  
CURRENT APPLICATION NUMBER: US/10/001,885  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/252,061  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: 60/253,257  
PRIOR FILING DATE: 2000-11-27  
NUMBER OF SEQ ID NOS: 167  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 125  
LENGTH: 121  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-001-885-125  
Query Match 100.0%; Score 632; DB 12; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2.4e-58;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVRILANGELIVQDDPRVRRITTOPPRGSIPIROSFFNRHGAPPGGPGRQOAGARIGAA 60  
DB 1 MVRILANGELIVQDDPRVRRITTOPPRGSIPIROSFFNRHGAPPGGPGRQOAGARIGAA 60  
QY 61 OSPFDLRLQVWNGFPGWHGNGAVPEPVISILLFILMTGVGGLLVGLVIVSHLSQ 120  
DB 61 OSPFDLRLQVWNGFPGWHGNGAVPEPVISILLFILMTGVGGLLVGLVIVSHLSQ 120

QY 121 R 121  
DB 121 R 121

## RESULT 2

US-10-076-747-122  
; Sequence 122, Application US/10076747  
; Publication No. US20030180725A1  
; GENERAL INFORMATION:  
; APPLICANT: Salceda, Susana  
; APPLICANT: Macina, Roberto  
; APPLICANT: Hu, Ping  
; APPLICANT: Recipon, Heve  
; APPLICANT: Katta, Kalpana  
; APPLICANT: Caffery, Robert  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and F  
; FILE REFERENCE: DEX-0315  
; CURRENT APPLICATION NUMBER: US/10/076,747  
; PRIOR FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: 60/266,290  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/266,834  
; PRIOR FILING DATE: 2001-02-15  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 122  
; LENGTH: 121  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-076-747-122

## Query Match

Best Local Similarity 100.0%; Score 632; DB 14; Length 121;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGIEIVODDDPRVRYTTTQPPRGSIIPROSFNNGHGAPPGGPPROQOAGARLGAA 60  
DB 1 MVRILANGIEIVODDDPRVRYTTTQPPRGSIIPROSFNNGHGAPPGGPPROQOAGARLGAA 60  
QY 61 QSPFDLNRQLVNGFPGMHGNAVEPVTSIILLFILMLGVYGLLLGVYVLSHSQ 120  
DB 61 QSPFDLNRQLVNGFPGMHGNAVEPVTSIILLFILMLGVYGLLLGVYVLSHSQ 120  
QY 121 R 121  
DB 121 R 121

## RESULT 3

US-09-809-391-673  
; Sequence 673, Application US/09809391  
; Publication No. US20030049618A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/809,391  
; PRIOR FILING DATE: 2001-03-16  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 761  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 673  
; LENGTH: 111  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-809-391-673

Query Match 61.2%; Score 387; DB 10; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1e-32;

Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGIEIVODDDPRVRYTTTQPPRGSIIPROSFNNGHGAPPGGPPROQOAGARLGAA 60  
DB 39 MVRILANGIEIVODDDPRVRYTTTQPPRGSIIPROSFNNGHGAPPGGPPROQOAGARLGAA 98

QY 61 QSPFDLNRQLVNGFPGMHGNAVEPVTSIILLFILMLGVYGLLLGVYVLSHSQ 120  
DB 99 QSPFDLNRQLVNGFPGMHGNAVEPVTSIILLFILMLGVYGLLLGVYVLSHSQ 120

## RESULT 4

US-09-882-171-673  
; Sequence 673, Application US/09882171  
; Publication No. US20030175858A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/882,171  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/809,391  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/149,476  
; PRIOR FILING DATE: 1998-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/04493  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: 60/040,162  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,333  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/038,621  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,626  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,334  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,336  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,163  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/047,600  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,615  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,597  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,502  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,633  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,583  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,617  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,618  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,503  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,592  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,581  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,584  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,500  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,587  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,492  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,598  
; PRIOR FILING DATE: 1997-05-23

;; PRIOR APPLICATION NUMBER: 60/047,613  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,582  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,596  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,612  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,632  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,601  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,580  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,568  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,314  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,569  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,311  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,671  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,674  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,669  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,312  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,313  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,672  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,315  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/048,974  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/056,886  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,877  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,889  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,893  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,630  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,878  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,662  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,872  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,882  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,637  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,903  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,888  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,879  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,880  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,894  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,911  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,636  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,874

;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,910  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,864  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,631  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,845  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,892  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/057,761  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/047,595  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,599  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,590  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,588  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,585  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,586  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,590  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,594  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,589  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,593  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,614  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,578  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,576  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/047,501  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,670  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/056,632  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,664  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,876  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,881  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,909  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,875  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,862  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,887  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,908  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/048,964  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/057,650  
;; PRIOR FILING DATE: 1997-09-05  
;; PRIOR APPLICATION NUMBER: 60/056,884  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/057,669  
;; PRIOR FILING DATE: 1997-09-05

Query Match 61.2%; Score 387; DB 10; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1e-32;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 WRIILANGBIYDDDDPRVATTTQPPRGSIPIROSFTNRGHGAPGCGPPROCGAGRLGA 60

Tue Aug 24 09:48:30 2004

us-10-001-885-125.rapb

Page 4

DB 39 MVRILANGEIYODDDPRVTTTOPPRGSIIPROSEFNGHGAAPGPGPPOQAAGARLGAA 98  
QY 61 QSPFNDLNRQLYN 73  
DB 99 QSPFNDLNRQLYN 111

RESULT 5  
US-10-164-861-673  
; Sequence 673, Application US/10164861  
; Publication No. US20030225248A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1  
; CURRENT APPLICATION NUMBER: US/10/164,861  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: US/09/149,476  
; PRIOR FILING DATE: 1998-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/04493  
; PRIOR FILING DATE: 1998-03-06  
; NUMBER OF SEQ ID NOS: 757  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 673  
; LENGTH: 111  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-164-861-673

Query Match 61.2%; Score 387; DB 12; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1e-32;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGEIYODDDPRVTTTOPPRGSIIPROSEFNGHGAAPGPGPPOQAAGARLGAA 60  
DB 39 MVRILANGEIYODDDPRVTTTOPPRGSIIPROSEFNGHGAAPGPGPPOQAAGARLGAA 98  
QY 61 QSPFNDLNRQLYN 73  
DB 99 QSPFNDLNRQLYN 111

RESULT 6  
US-09-809-391-367  
; Sequence 367, Application US/09809391  
; Publication No. US20030049618A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/809,391  
; CURRENT FILING DATE: 2001-03-16  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 761  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 367  
; LENGTH: 48  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-809-391-367

Query Match 38.8%; Score 245; DB 10; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2.9e-18;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 MGFPQWHLGNHAEVPTISILLFILMLGVRGLLVGLVYVSHLSOR 121  
DB 1 MGFPQWHLGNHAEVPTISILLFILMLGVRGLLVGLVYVSHLSOR 48

RESULT 7

; Sequence 367, Application US/09882171  
; Publication No. US20030175858A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 166 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/882,171  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 09/809,391  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/149,476  
; PRIOR FILING DATE: 1998-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/04493  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: 60/040,162  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,333  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/038,621  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,626  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,334  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,336  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/040,163  
; PRIOR FILING DATE: 1997-03-07  
; PRIOR APPLICATION NUMBER: 60/047,600  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,615  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,597  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,502  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,633  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,583  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,617  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,618  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,503  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,592  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,581  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,584  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,500  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,587  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,492  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,598  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,613  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,582  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,596  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,612  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,632  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/047,601  
; PRIOR FILING DATE: 1997-05-23

;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,568  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,314  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,569  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,311  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,671  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,674  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,669  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,312  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,313  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,672  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,315  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/048,974  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/056,886  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,877  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,889  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,893  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,630  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,878  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,662  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,872  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,882  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,637  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,903  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,888  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,879  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,880  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,894  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,911  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,636  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,874  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,910  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,864  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,631  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,845  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,892  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/057,761  
;; PRIOR FILING DATE: 1997-08-22

;; PRIOR APPLICATION NUMBER: 60/047,595  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,599  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,588  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,585  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,586  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,590  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,594  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,589  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,593  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,614  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,578  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,576  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/047,501  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,670  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/056,632  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,881  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,909  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,875  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,862  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,887  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,908  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/048,964  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/057,650  
;; PRIOR FILING DATE: 1997-09-05  
;; PRIOR APPLICATION NUMBER: 60/056,884  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/057,669  
;; PRIOR FILING DATE: 1997-09-05

Query Match 38.8%; Score 245; DB 10; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2,9e-18;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 MGPQWHLGNHAEVPTSTILLFLMTGVRGLLVGLVYVSHLSQR 121  
Db 1 MGPQWHLGNHAEVPTSTILLFLMTGVRGLLVGLVYVSHLSQR 48

RESULT 8  
US-10-164-861-367  
; Sequence 367, Application US/10164861  
; Publication No. US20030225248A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1  
; CURRENT APPLICATION NUMBER: US/10/164,861

Tue Aug 24 09:48:30 2004

us-10-001-885-125.rapb

Page 6

```

; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US/09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 757
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 367
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-861-367

```

```

Query Match 38.8%; Score 245; DB 12; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.9e-18;
Matches 48; Conservative 0; Mismatches 0; Gaps 0;

```

```

QY 74 MGFPQWHLGNHVEPTSTILLFLMLGVKGLLVGLVYVSHLSQR 121
DB 1 MGFPQWHLGNHVEPTSTILLFLMLGVKGLLVGLVYVSHLSQR 48

```

```

RESULT 9
US-10-108-260A-3882
; Sequence 3882, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3882
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3882

```

```

Query Match 17.5%; Score 110.5; DB 15; Length 132;
Best Local Similarity 31.9%; Pred. No. 0.0012;
Matches 30; Conservative 17; Mismatches 34; Indels 13; Gaps 3;

```

```

QY 26 RGSIPRQS-----FNRGHGAPPGGPRQOQAGARLGAASPFNDLNRLQVNMGFPQWHL 81
DB 42 RGRPKESQGVDSQNTGEPVGDYK-----KMGTL--FGLKNKLNMGFTMYF 92
QY 82 GNHVEPTSTILLFLMLGVKGLLVGLVYV 115
DB 93 GERIVEPTIVTFWVLMFLGLQALGLVAVLCIV 126

```

```

RESULT 10
US-10-156-761-10980
; Sequence 10980, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697

```

```

; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10980
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10980

```

```

Query Match 15.7%; Score 99.5; DB 14; Length 223;
Best Local Similarity 30.9%; Pred. No. 0.031; Indels 35; Gaps 4;
Matches 29; Conservative 9; Mismatches 21;

```

```

QY 23 QPP-----RGSIPRQS-----FNRGHGAP--PGGPRQOQAGARLGAASPFNDLNRL 71
DB 27 QPPYQAPGRIPOGQGVGYPGGPGGPRVAMGRRLLGA----- 74
QY 72 VNMGFPQWHLGNHVEPTSTILLFLMLGVKGLLVGLVYVSHLSQR 105
DB 75 -----RAIDGVAFFVTFILGALGVAG 96

```

```

RESULT 11
US-10-424-599-174474
; Sequence 174474, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174474
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_WRT3847_128569C.1.pep
US-10-424-599-174474

```

```

Query Match 13.5%; Score 85.5; DB 12; Length 112;
Best Local Similarity 32.8%; Pred. No. 0.4; Indels 11; Gaps 2;
Matches 22; Conservative 8; Mismatches 26;

```

```

QY 37 RGHGAPPGGPRQOQAG-ARLGAASPFND-----LNRLQVNMGFPQWHLGNH 85
DB 5 RGGSPRPGKGRSAAKGRAGAKPPNPULVFPAGVILLNKQPPFALPMKIPVSK 64
QY 86 VEPVTSI 92
DB 65 SEPIAI 71

```

```

RESULT 12
US-10-437-963-198292
; Sequence 198292, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

```

```

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198292
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_93967C.1.pep
US-10-437-963-198292

Query Match 13.0%; Score 82; DB 16; Length 123;
Best Local Similarity 42.9%; Pred. No. 1;
Matches 27; Conservative 4; Mismatches 14; Indels 18; Gaps 5;

QY 17 RVRTTTPPGSGIPRSGF-FNRGHGAP-----PGGPGPQQGAGAR----LGAA 60
DB 25 RGRTRT-PTRGSPRSAPFPRRGPGPPRQGLRAGAGGP-PPGRRGARGAPLHVAG 82

QY 61 QSP 63
DB 83 KSP 85

RESULT 13
US-10-425-114-60982
; Sequence 60982, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60982
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-039-A12_F11.pep
US-10-425-114-60982

Query Match 12.8%; Score 81; DB 12; Length 256;
Best Local Similarity 33.0%; Pred. No. 3.1;
Matches 29; Conservative 7; Mismatches 26; Indels 26; Gaps 4;

QY 26 RGSIPRGSFNRGHGAPPGGPGP-----RQQAARIGA-----AQSPNDLNR 69
DB 81 RGAUFR--LIRGGAGQGVPPAGRRGLRRRRRRRAGARPGAHPRPAOVHGRRHPLRLH 137

QY 70 QLVNM-----GPPQHLGNHAVEPYT 90
DB 138 RHVRRALRHENGDPPDERGVHAAVPT 165

RESULT 14
US-10-106-698-8371
; Sequence 8371, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005F1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
```

```

; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 8371
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (88)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (91)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-8371

Query Match 12.7%; Score 80; DB 14; Length 92;
Best Local Similarity 35.7%; Pred. No. 1.2;
Matches 30; Conservative 2; Mismatches 26; Indels 26; Gaps 4;

QY 15 DPRVR-----TTTPRGSIPRSGFNRGHGAPPGGPGPQQGAGARGAASPFND 66
DB 15 DPRVPRSSSGSPLYTRVPRAS-PAPEGPSGHALPGAGLR--AGHERGAAR---- 66

QY 67 LNRQLVMGFPQWHLGNHAVEPYT 90
DB 67 -----PRAHGGRSAARGT 80

RESULT 15
US-10-156-761-10160
; Sequence 10160, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OKURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10160
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10160

Query Match 12.7%; Score 80; DB 14; Length 392;
Best Local Similarity 27.2%; Pred. No. 6.6;
Matches 37; Conservative 18; Mismatches 33; Indels 48; Gaps 8;

QY 7 NGEIYQDDDPRT-----TTTPRGSIPRSGFNRGHG-----APPGGPGP 48
```

Tue Aug 24 09:48:30 2004

us-10-001-885-125.rapb

Page 8

```
Db 67 SGPLFRDDTPQTTSYGSQTPYVTQSFPYGS-----QTPHGAOPQKPEEMPDAPQPAFAP 119
OY 49 ROOQAGARLGAAGSPFNDLNRQLVNMGFPOWHLGNHAEVETSILLFL--NMLGVRGL 106
Db 120 QKSNAGRDLAGAIG-----VGVG-----LG-----VVIWASLFFVKAVFVGVIAY 159
OY 107 -LIVGLVTVLSHSOR 121
Db 160 AVVVGILWELTSRIER 175
```

Search completed: August 24, 2004, 09:33:21  
Job time : 126 secs



OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2004, 07:11:30 ; Search time 2048 Seconds  
(without alignments)  
2560.794 Million cell updates/sec

Title: US-10-001-885-125  
Perfect score: 632  
Sequence: 1 MVTILNGELVQDDPRVRT.....GVKGLLVGVLYVLSLSQR 121

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xld  
-O=/cgnr/1/USPTO.spool/p/US10001885/runat/17082004.151718.25904/app\_query\_faeta\_1.263  
-DB=GenMdb1 -QMT=Fastcap -SUFFIX=p2n.ige -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=Dblosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10001885 @CGN 1.1 5600 @runat/17082004.151718.25904 -NCPU=6 -ICPU=3  
-NO MAP -IARGOUTDIR -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAXN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : GenMdb1:\*  
1: gb ba:\*  
2: gb htg:\*  
3: gb ln:\*  
4: gb om:\*  
5: gb ov:\*  
6: gb pat:\*  
7: gb ph:\*  
8: gb pl:\*  
9: gb pr:\*  
10: gb ro:\*  
11: gb sts:\*  
12: gb sy:\*  
13: gb un:\*  
14: gb vl:\*  
15: em ba:\*  
16: em fun:\*  
17: em hum:\*  
18: em ln:\*  
19: em mu:\*  
20: em om:\*  
21: em or:\*  
22: em ov:\*  
23: em pac:\*  
24: em ph:\*  
25: em pl:\*  
26: em ro:\*  
27: em sts:\*  
28: em un:\*

29: em vl:\*  
30: em htg hum:\*  
31: em htg inv:\*  
32: em htg other:\*  
33: em htg mus:\*  
34: em htg pin:\*  
35: em htg rod:\*  
36: em htg mam:\*  
37: em htg yrt:\*  
38: em sy:\*  
39: em htgo hum:\*  
40: em htgo mus:\*  
41: em htgo other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID       | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 632   | 100.0       | 1049   | 6 BD191093  | BD191093 186 human |
| 2          | 632   | 100.0       | 1129   | 6 BC013587  | BC013587 Homo sapi |
| 3          | 632   | 100.0       | 1192   | 6 AX431088  | AX431088 Sequence  |
| 4          | 577   | 91.3        | 550    | 6 AX113082  | AX113082 Sequence  |
| 5          | 568.5 | 90.0        | 1538   | 10 BC024943 | BC024943 Mus muscu |
| 6          | 563.5 | 89.2        | 275924 | 2 AC116233  | AC116233 Rattus no |
| 7          | 551   | 87.2        | 483    | 6 AR413227  | AR413227 Sequence  |
| 8          | 551   | 87.2        | 483    | 6 BD108780  | BD108780 EST and e |
| 9          | 544.5 | 86.2        | 240657 | 2 AC111675  | AC111675 Rattus no |
| 10         | 503   | 79.6        | 519    | 6 AR415427  | AR415427 Sequence  |
| 11         | 503   | 79.6        | 519    | 6 BD110980  | BD110980 EST and e |
| 12         | 472   | 74.7        | 165110 | 9 AL450311  | AL450311 Human DNA |
| 13         | 472   | 74.7        | 173341 | 2 AC021954  | AC021954 Homo sapi |
| 14         | 459   | 72.6        | 215050 | 2 AC127417  | AC127417 Mus muscu |
| 15         | 448   | 70.9        | 258815 | 2 AC127817  | AC127817 Rattus no |
| 16         | 443.5 | 70.2        | 138070 | 2 AC109783  | AC109783 Mus muscu |
| 17         | 135   | 21.4        | 754    | 3 AK113891  | AK113891 Clona int |
| 18         | 130   | 20.6        | 238937 | 2 AC094926  | AC094926 Rattus no |
| 19         | 110.5 | 17.5        | 2594   | 6 AR834315  | AR834315 Sequence  |
| 20         | 110.5 | 17.5        | 2594   | 9 AC096689  | AC096689 Homo sapi |
| 21         | 110   | 17.4        | 2565   | 9 BC022534  | BC022534 Homo sapi |
| 22         | 107.5 | 17.0        | 109668 | 2 AC109347  | AC109347 Homo sapi |
| 23         | 107.5 | 17.0        | 170662 | 2 BX005193  | BX005193 Danto rer |
| 24         | 107.5 | 17.0        | 172287 | 2 AC069028  | AC069028 Homo sapi |
| 25         | 107   | 16.9        | 155320 | 2 AC135762  | AC135762 Rattus no |
| 26         | 107   | 16.9        | 241394 | 2 AC131549  | AC131549 Rattus no |
| 27         | 105   | 16.6        | 496    | 10 BC039568 | BC039568 Mus muscu |
| 28         | 105   | 16.6        | 151241 | 2 AC115041  | AC115041 Mus muscu |
| 29         | 105   | 16.6        | 153126 | 2 AC102465  | AC102465 Mus muscu |
| 30         | 101   | 16.0        | 1383   | 6 AR85066   | AR85066 Sequence   |
| 31         | 101   | 16.0        | 212840 | 2 AC135223  | AC135223 Gallus ga |
| 32         | 99.5  | 15.7        | 183599 | 2 AC073501  | AC073501 Homo sapi |
| 33         | 99.5  | 15.7        | 296300 | 1 AP005035  | AP005035 Streptomy |
| 34         | 98.5  | 15.6        | 149610 | 8 AP003244  | AP003244 Oryza sat |
| 35         | 98.5  | 15.6        | 163513 | 8 AP003200  | AP003200 Oryza sat |
| 36         | 96.5  | 15.3        | 66254  | 9 AL137791  | AL137791 Human DNA |
| 37         | 96.5  | 15.3        | 179221 | 9 AC114493  | AC114493 Homo sapi |
| 38         | 95    | 15.0        | 3705   | 5 CG853435  | CG853435 Gallus ga |
| 39         | 94.5  | 15.0        | 10348  | 1 AB004846  | AB004846 Pseudomn  |
| 40         | 94.5  | 15.0        | 177226 | 1 AC121858  | AC121858 Mus muscu |
| 41         | 94.5  | 15.0        | 183975 | 10 AL590988 | AL590988 Mouse DNA |
| 42         | 94.5  | 15.0        | 248358 | 2 AC122944  | AC122944 Rattus no |
| 43         | 93.5  | 14.8        | 156997 | 9 AC011465  | AC011465 Homo sapi |
| 44         | 93    | 14.7        | 236639 | 2 AC127406  | AC127406 Rattus no |
| 45         | 93    | 14.7        | 254489 | 2 AC098275  | AC098275 Rattus no |

RESULT 1

ALIGNMENTS

BD191093 1049 bp DNA linear PAT 17-JUL-2003  
 LOCUS 186 human secreted proteins.  
 DEFINITION BD191093.1 GI:33000832  
 VERSION JP 2002510192-A/57.  
 KEYWORDS unidentified  
 SOURCE unidentified  
 ORGANISM unidentified

REFERENCE  
 AUTHORS 1 (bases 1 to 1049)  
 Ruben,S.M., Rosen,C.A., Fischer,C.L., Soppet,D.R., Carter,K.C.,  
 Bednarik,D.P., Endress,G.A., Yu,G.L., N.J., Feng,P., Young,P.E.,  
 Greene,J.M., Ferris,A.M., Duan,R., Hu,J.S., Florence,K.A.,  
 Olsen,H.S., Ebner,R., Brewer,L.A., Moore,P.A., Shi,Y.,  
 Lafleur,D.W., Li,Y., Zeng,Z. and Kyaw,H.

TITLE  
 JOURNAL 186 human secreted proteins  
 Patent: JP 2002510192-A 57 02-APR-2002;  
 HUMAN GENOME SCIENCES INC

COMMENT  
 PN JP 2002510192-A/57  
 PD 02-APR-2002  
 PF 06-MAR-1998 JP 1998538883  
 PR 07-MAR-1997 US 60/040162, 07-MAR-1997 US 60/040333 PR  
 07-MAR-1997 US 60/038621, 07-MAR-1997 US 60/040161 PR  
 07-MAR-1997 US 60/040626, 07-MAR-1997 US 60/040334 PR  
 07-MAR-1997 US 60/040336, 07-MAR-1997 US 60/040163 PR  
 11-APR-1997 US 60/043580, 11-APR-1997 US 60/043568 PI  
 M RUBEN, CRAIG A ROSEN, CARIE L FISCHER, DANIEL R SOPPET, PI  
 KENNETH C CARTER, DANIEL P BEDNARIK, GREGORY  
 A ENDRESS, GUO LIANG  
 PI YU, JIAN NI  
 PI PING FENG, PAUL E YOUNG, JOHN M GREENE, ANN  
 M FERRIE, ROXANNE DUAN,  
 PI JING SHAN HU, KIMBERLY A FLORENCE, HENRIK  
 S OLSEN, REINHARD EBNER,  
 PI LAURIE A BREWER, PAUL A MOORE, YANGGU SHI, DAVID W LAFLEUR PI  
 YI LI, ZHI ZHEN ZENG,  
 PI HUA KYAW  
 PC C12N15/12, C12N5/10, C12N1/21, C07K14/47, C07K16/18, C12Q1/68, PC  
 G01N33/50,  
 PC G01N33/53, G01N33/68, A61K38/17  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 FH Key

FEATURES  
 source Location/Qualifiers.  
 1..1049 Location/Qualifiers  
 1..1049 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

ORIGIN

Alignment Scores:  
 Pred. No.: 1,52e-43 Length: 1049  
 Score: 632.00 Matches: 121  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-001-885-125 (1-121) x BD191093 (1-1049)

QY 1 MetValaTgIleLeuValaAsnGlyGluIleValaGlnAspAspProArgValaTgThr 20  
 Db 116 ATGCGCGGATCTTGCCCAATGGGGAATCGTGCAGCGACGACCCCGAGTGAACACC 175  
 QY 21 ThrThnGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyVH:Gly 40  
 Db 176 ACTACCCAGCCACCAAGAGTGAATCTCTGACAGAGACTTCTTCATTAAGGAGCCATGGT 235  
 QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnaGlyValaArgLeuGlyValaAla 60  
 Db 236 GTTCCCCAGAGGGGCTCTGGCCCCCGCCAGACGACGAGGCGTGCAGGCTGGTGTCT 295

Db 296 CAGTCCCTCCCTCAATGACCTCAACCGGACGCTGGGAACAGGGGCTTCCGACAGTGCAT 355  
 QY 81 LeuGlyAamHialaValaGluProValaThrSerIleLeuLeuPheLeuMetMet 100  
 Db 356 CTGGGACACATGCTGTGAGCGGTGACCTTCATCTCTCTCTCTCTCTCATATG 415  
 QY 101 LeuGlyValaArgGlyLeuLeuLeuValaGluValaIlyLeuValaSerHisLeuSerGln 120  
 Db 416 CTGGGTGTCGTGGCTCTCTCTCTGTGGCTTGTCTACTGTCGTCACCTGATGTCAG 475  
 QY 121 Arg 121  
 Db 476 CGG 478

RESULT 2  
 BC013587 1129 bp mRNA linear PRI 11-DEC-2003  
 LOCUS Homo sapiens chromosome 10 open reading frame 35, mRNA (cdna clone  
 DEFINITION MGC:9596 IMAGE:389656), complete cds.  
 ACCESSION BC013587  
 VERSION BC013587.1 GI:15488919  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1129)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stacheton,M., Soares,M.B., Bonaldi,M.P., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,N.J., Uedlin,T.B., Tosilylki,S.,  
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Mallaby,S.J., Bosak,S.A., McSwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Holly,S.M.,  
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Raney,V., Helton,E., Kerteman,M., Madan,A., Rodriguez,S.,  
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butterfield,Y.S., Krzywinski,M.I., Skalka,U., Smalins,D.S.,  
 Scherch,A., Schein,J.B., Jones,S.J. and Marra,M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 1129)  
 Strausberg,R.  
 Direct Submission  
 Submitted (04-SEP-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. W.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILMIL at: <http://image.llnl.gov>

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21687152. Location/Qualifiers

#### FEATURES

source

1..1129  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="WGC:9596 IMAGE:389656"  
/tissue\_type="Pancreas, epithelioid carcinoma"  
/clone\_id="NH1 MG\_70"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..1129  
/gene="C10orf35"  
/db\_xref="LocusID:219738"  
160..525  
/codon\_start=1  
/product="chromosome 10 open reading frame 35"  
/protein\_id="AAH3587.1"  
/db\_xref="GI:15488920"  
/db\_xref="LocusID:219738"  
/translation="MVRILANGELIVQDDPRVTTQPPRGSIPIROSFENRGGAPPG  
GPRQOQAGARLGAQSPFNDLNRQLVNMGPQMHGNHVAPEVTISILLFPLMLG  
VRGILLVGLVIVLSHSQR"

#### ORIGIN

Alignment Scores:  
Pred. No.: 1.64e-43 Length: 1129  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-001-885-125 (1-121) x BC013587 (1-1129)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
Db 160 ATGTTGGGGAATCTTGGCCAAATGGGGAATGTGGAGATAGACACCCCGAGTGAAGACC 219  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHis 40  
Db 220 ACTAACCAAGCACCAGAGGTAGCATTCCTGACAGAGCTTCTCAATAGGGCCATGTGT 279  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnAlaGlyValAlaArgLeuGlyAla 60  
Db 280 GCTCCCAAGGGGCTCTGGCCCCCGCCAGCAGCAGGAGCTGGGTGCTGCT 339  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
Db 340 CAGTCCCTTCATGACCTCAACCGGACGCTGGTGAACATGGGCTTCCGAGTGGCAT 399  
QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
Db 400 CTCGGCAACATGCTGTGAGACCGGTGACCTCCATCTCTCTCTCTCTCATGATG 459  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
Db 460 CTTCGTGTGTGTGCTCTCTCTCTGTGGCTGTCTACTGATGATCCACCTGAGTCA 519  
QY 121 Arg 121  
Db 520 CGG 522

RESULT 3  
AX431088 1192 bp DNA linear PAT 28-JUN-2002  
LOCUS AX431088  
DEFINITION Sequence 31 from Patent WO0240535.  
ACCESSION AX431088  
VERSION AX431088.1 GI:21656065  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
1 Recipon,H., Liu,C., Salceda,S., Sun,Y., Cafferkey,R. and Macina,R.A.  
TITLE  
1 Compositions and methods relating to ovary specific genes and proteins  
JOURNAL  
1 Patent: WO 0240535-A 31 23-MAY-2002;  
DIADEXUS INC (US)  
FEATURES  
source  
Location/Qualifiers  
1..1192  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

#### ORIGIN

Alignment Scores:  
Pred. No.: 1.74e-43 Length: 1192  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-001-885-125 (1-121) x AX431088 (1-1192)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
Db 161 ATGTTGGGGAATCTTGGCCAAATCTGACAGACAGACCCCGAGTGAAGACC 220  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHis 40  
Db 221 ACTAACCAAGCACCAGAGGTAGCATTCCTGACAGAGCTTCTCAATAGGGCCATGTGT 280  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnAlaGlyValAlaArgLeuGlyAla 60  
Db 281 GCTCCCAAGGGGCTCTGGCCCCCGCCAGCAGCAGGAGCTGGGTGCTGCT 340  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
Db 341 CAGTCCCTTCATGACCTCAACCGGACGCTGGTGAACATGGGCTTCCGAGTGGCAT 400  
QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
Db 401 CTCGGCAACATGCTGTGAGACCGGTGACCTCCATCTCTCTCTCTCATGATG 460  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
Db 461 CTTCGTGTGTGTGCTCTCTCTGTGGCTGTCTACTGATGATCCACCTGAGTCA 520  
QY 121 Arg 121  
Db 521 CGG 523

RESULT 4  
AX113082 550 bp DNA linear PAT 01-MAY-2001  
LOCUS AX113082  
DEFINITION Sequence 5 from Patent WO0127276.  
ACCESSION AX113082  
VERSION AX113082.1 GI:13939514  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
1 Harlocker,S.L., Dillon,D.C. and Xu,T.  
TITLE  
1 DNA sequences from breast tumor and uses thereof  
JOURNAL  
1 Patent: WO 0127276-A 5 19-APR-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source  
Location/Qualifiers  
1..550  
/organism="Homo sapiens"

misc\_feature /mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..550  
/note="n = A,T,C or G"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,76e-39 Length: 550  
Score: 577.00 Matches: 117  
Percent Similarity: 97.52% Conservative: 1  
Best Local Similarity: 96.69% Mismatches: 2  
Query Match: 91.30% Indels: 2  
DB: Gaps: 0

US-10-001-885-125 (1-121) x AX113082 (1-550)

CY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 115 ATGTGGCGGATCTGGCCCATGCGAAATCTGCAGACACACACCCCGACGAGACC 174  
CY 21 ThrThgInpProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40  
DB 175 ACTACCGACGACCAAGAGGTAGCATTCCTCGACAGAGCTTTCATATAGGGCCCATGTGT 234  
CY 41 AlaProPArgGlyGlyProGlyProArgGlnGlnIleValGlnAlaArgGlyValAla 60  
DB 235 GCTCCCCCGGGGGGCTGCGCCCGCCGACGAGGAGGTGCCAGGCTGGTGCTGT 294  
CY 61 GlnSerProPheAsnArgGlyLeuValAsnMetGlyPheProGlnITPHis 80  
DB 295 CAGTCCCCCTTCATGACCTCAACCGGAGCTGTGAACATGGGCTTCCGCAATGGCAT 354  
CY 81 LeuGlyAsnHisAla-ValGluProValThrSerIleLeuLeuPheLeuMetMet 100  
DB 355 CTGGGCAACCATGCTGTGGAGCGCGGTACCTCATCTCTGT-CTCTTCTGCTCATGAT 413  
CY 100 tLeuGlyValArgGlyLeuLeuValGlyLeuValIleValIleValSerHisLeuSerG 120  
DB 414 GCTTGCTGTGCTGTGCTTCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 473  
CY 120 n 120  
DB 474 A 474

RESULT 5  
BC024943 1538 bp mRNA linear ROD 06-OCT-2003  
LOCUS Mus musculus RIKEN cDNA 2010107G23 gene, mRNA (cDNA clone MGC:28201  
DEFINITION IMAGE:3989515), complete cds.  
ACCESSION BC024943  
VERSION BC024943.1 GI:19354288  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1538)  
Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner K.H., Schaeffer C.F., Bhac N.K.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaeffer C.F., Bhac N.K.,  
Hopkins R.F., Jordan H., Moore T., Nak S.I., Wang J., Heideh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stableton M., Soares M.B., Bonaldo M.F., Casavant T.L.,  
Schetz T.E., Brownstein M.J., Uesdin T.B., Toshikiyuki S.,  
Carroll P., Prange C., Raha S.S., Loughlino N.A., Peters G.J.,  
Abramsen R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,  
McKernan K.J., Malek J.A., Gamarale P.H., Richards S.,  
Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton B., Kettelman M., Madan A.C., Rodriguez S.,  
Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,  
Bonfatti G.G., Blakeley R.W., Touchman J.W., Green B.D.

Butterfield, Y.S., Krzywinaki, M.I., Skalska, U., Smalhus, D.E.,  
Schmerch, A., Schein, J.E., Jones, S.J., and Marr, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
JOURNAL MEDLINE  
22388257  
PUBMED 12477932  
TITLE 2 (bases 1 to 1538)  
AUTHORS Strusberg, R.  
REFERENCE Direct Submission  
Submitted (01-MAR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gamarale, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Louisse, H.,  
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nambavali,  
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/BLNI at: <http://image.llnl.gov>  
Series: IRAX Project: 36 Row: h Column: 13  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein.

## FEATURES

source

1..1538  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CECH 11"  
/db\_xref="taxon:10090"  
/clone="MGC:28201 IMAGE:3989515"  
/tissue="mammary tumor metastasized to lung. Tumor  
arose spontaneously from a senescent normal mammary  
(clonal) outgrowth infected with the virus MMV."  
/clone\_1lb="NCI CGAP\_Lu29"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"  
1..1538  
/gene="2010107G23R1k"  
/db\_xref="LOCUSID:69894"  
/db\_xref="MGI:1917144"  
486..848  
/codon\_start=1  
/product="2010107G23R1k protein"  
/protein\_id="AAH24943.1"  
/db\_xref="GI:19354289"  
/db\_xref="GI:19354289"  
/translation="MRLIANGELIVQDDPFRVTTTQRRSSQCGFRNRHGAPPG  
PSPROQAGARLGAAGSPSILNRQLVNMGPQWHLGNHVEPVTSILLPLFMKVLGV  
RGLIIVGLVIVSHISQR"

## CDS

ORIGIN  
Alignment Scores:  
Pred. No.: 4.08e-38 Length: 1538  
Score: 568.50 Matches: 112  
Percent Similarity: 94.21% Conservative: 2  
Best Local Similarity: 92.56% Mismatches: 6  
Query Match: 89.95% Indels: 1  
DB: Gaps: 1

US-10-001-885-125 (1-121) x BC024943 (1-1538)



FEATURES \* 274719 275924: contig of 1206 bp in length.  
Location/Qualifiers  
1..275924  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-95L17"  
2610..3635  
/note="wgs end extension  
clone end:Sp6"  
misc\_feature  
9666..10838  
/note="wgs end extension  
clone end:Sp6"  
misc\_feature  
10939..12161  
/note="wgs end extension  
clone end:Sp6"  
misc\_feature  
91900..93203  
/note="wgs end extension  
clone end:Sp6"  
misc\_feature  
135055..135316  
/note="clone boundary  
clone end:Sp6  
site:EcoRI  
end sequence: BH361167"  
165565..210774  
/note="clone boundary  
clone end:T7  
site:EcoRI  
end sequence: BH361166"  
misc\_feature  
ALIGNMENT SCORES:  
Pred. No.: 2,386-35 Length: 275924  
Score: 563.50 Matches: 111  
Percent Similarity: 93.39% Conservative: 2  
Best Local Similarity: 91.74% Mismatches: 7  
Query Match: 89.16% Indels: 1  
DB: 2 Gaps: 1  
US-10-001-885-125 (1-121) x AC116233 (1-275924)  
QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 227956 ATGTGAGGATCTTGCTAATGGGAGATCGTTCAAGATGATGACCCCGAGTGAAGCAT 227897  
QY 21 ThrThgInProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 227896 ACCACCCCA---CAGAGGAGTGAAGCAGCCGCGAGGAGGTGCGGCTTTTTCACAGAGCCACAGT 227840  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 227839 GCAACCTCCAGGGGGGCGTGAAGCCCGCAGCAGCAGGAGTGGCCGACTGGGTGCTGCC 227780  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 227779 CAATCTCCCTTCAAGTACCTGAACCGGAGCGTGTGAACATGGGCTTCCACCAATGGCAT 227720  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
DB 227719 CTTGGGAACCACTGCTGAGCTGTGACCTCACTCCCTGCTCTTCTCTCATGATA 227660  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValIleLeuValSerHisLeuSerGln 120  
DB 227659 CTGGGGGTTCTGTGGCTCTCTGCTGTGGGTCTGTGCTCACTGGTGTCTCACCTGAGCCAG 227600  
QY 121 Arg 121  
DB 227599 CGG 227597  
RESULT 7  
AR413227  
LOCUS 483 bp DNA linear PAT 18-DEC-2003

ACCESSION AR413227  
VERSION AR413227.1 GI:40168337  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 483)  
AUTHORS Edwards, J.-B. D. M., Jobert, S. and Giordano, J.-Y.  
TITLE EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 864 28-OCT-2003;  
FEATURES  
source  
1..483  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
ALIGNMENT SCORES:  
Pred. No.: 3,42e-37 Length: 483  
Score: 551.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 87.18% Indels: 0  
DB: 6 Gaps: 0  
US-10-001-885-125 (1-121) x AR413227 (1-483)  
QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 171 ATGTGAGGATCTTGCTAATGGGAGATCGTTCAAGATGATGACCCCGAGTGAAGCAT 230  
QY 21 ThrThgInProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 231 ACTACCCAGCCACCAAGAGTGAAGTCTCCGACGAGAGTCTTCAATAGGGGCGCATGCT 290  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 291 GCTCCCCCAGGGGTCCTGAGCCCGCAGCAGCAGGAGGAGTGGCTGCTGCT 350  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 351 CAGTCCCTTCAATGACCTCAACCGGAGCGTGTGAACATGGGCTTCCGACGTGGCAT 410  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
DB 411 CTGGGAACCACTGCTGAGCGCGGAGTCACTCTCTCTCTCTCATGATG 470  
QY 101 LeuGlyValArg 104  
DB 471 CTGGTGTCTGT 482  
RESULT 8  
BD108780 483 bp DNA linear PAT 18-SEP-2002  
BD108780  
DEFINITION  
EST and encoded human protein.  
ACCESSION  
BD108780.1 GI:23203598  
VERSION  
JP 2002010789-A/857.  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 (bases 1 to 483)  
AUTHORS Edwards, J. B. D. M., Jobert, S. and Giordano, J. E.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 857 15-JAN-2002;  
COMMENT  
GENSET CORP  
OS Homo sapiens (human)  
PN JP 2002010789-A/857  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 2000280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELLIN JOBERT, JEAN EYE PI

PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N15/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC EST and encoded human protein  
FH Key Location/Qualifiers  
FT CDS 171..482.  
1.483  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

Alignment Scores:  
Pred. No.: 3,42e-37 Length: 483  
Score: 551.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 87.18% Indels: 0  
Gaps: 0

US-10-001-885-125 (1-121) x BD108780 (1-483)  
QY 1 MetValArgIleLeuAlaAsnGlyValIleValGlnAspAspAppProArgValArgThr 20  
DB 171 ATGGCGCGGATCTTGCCCAATGGGAATCTGTGACGACGACGACCCCGAGTGAAGGCC 230  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyValIleGly 40  
DB 231 ACTACCCGACCAACCAAGAGTGAAGTCTCTGACAGAGCTTCTCAATAGGCGCCATGCT 290  
QY 41 AlaProProGlyValProGlyProArgGlnGlnGlnAlaGlyValArgLeuGlyValAla 60  
DB 291 GCTCCCGACGAGGGGCTCTGGCCCGCCGACGACGAGGAGTGCAGGCTGGGCTGCT 350  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnIleValAlaMetGlyPheProGlnTrpHis 80  
DB 351 CAGTCCCTTCAATGACCTCAACCGGACGCTGGTGAACAGGGCTTTCGCATGGCAT 410  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuPheLeuMetMet 100  
DB 411 CTCGGCAACCAATGCTGGAGACCGGCTGACCTTCCTGCTCTTCCTGCTCATGATG 470  
QY 101 LeuGlyValArg 104  
DB 471 CTGTGTTCT 482

RESULT 9  
AC11675/c 240657 bp DNA linear HTG 13-MAY-2003  
LOCUS Rattus norvegicus clone CH230-48G14, WORKING DRAFT SEQUENCE, 4  
DEFINITION  
AC11675.5 GI:30579217  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULFILL  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 240657)  
Muzny, D., Maric, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalbech, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryan, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyne, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K.,

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gerdegeorgis, E., Geier, K., Gilly, R., Grady, M., Guerra, W., Guevara, W.,  
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladik, S.L., Hodgson, A., Hogues, M.,  
Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C.L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, U., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorenz, H., Louie, H., Lozano, R.J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenan, E.,  
Milosavljevic, A., Miner, G., Minna, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Muniz, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwankwelu, O., Okunolu, G., Olapadesogun, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,  
Plummer, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,  
Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,  
Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
Shetty, V., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D.,  
Snead, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Willson, R., Wleczky, R., Woodson, H., Worely, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, S., Dunn, D., von  
Neidhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
Weinstock, G. and Gibbs, R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 240657)  
Worely, K.C.  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 240657)  
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 13, 2003 this sequence version replaced gi:24818892.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold, and  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GWMT  
Center clone name: CH230-48G14







KEYWORDS JP 2002010789-A/3057.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
AUTHORS 1 (bases 1 to 519)  
EDWARDS, J.B.D.M., JOBERT, S. and GIORDANO, J.E.  
TITLE EST and encoded human protein  
JOURNAL Parent: JP 2002010789-A 3057 15-JAN-2002;  
GENSET CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002010789-A/3057  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 2002080989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC EST and encoded human protein  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
1..519  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.46e-33 Length: 519  
Score: 503.00 Matches: 99  
Percent Similarity: 99.00% Conservatave: 0  
Best Local Similarity: 99.00% Mismatches: 1  
Query Match: 79.59% Indels: 1  
DB: Gaps: 0  
US-10-001-885-125 (1-121) x BD110980 (1-519)  
QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
Db 219 ATGTGTGGGATCTTGGCCAAATGGGAAATCGTCAGATATACGACCCCGAGTGAGACC 278  
QY 21 ThrThrglnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyIleGly 40  
Db 279 ACTACCCAGCCACAGAGGTAGCATTCCTCGACAGAGCTTCTCATATAGGGGCATGCT 338  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
Db 339 GCTCCCCCAGG-GGTCTGCGCCCGCCGACGACGACGACGAGGTGCMAGCTGCTGCTGCT 397  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValaAsnMetGlyPheProGlnTrpHis 80  
Db 398 CAGTCCCCCTTCATGACCTCAACCGGACGCTGATGACATGGGCTTCCGACATGGCAT 457  
QY 81 LeuGlyAsnHisAlaValaGlnProValThrSerIleLeuLeuLeuPheLeuLeuMetMet 100  
Db 458 CTCGGCAACCATCTGTGAGAGCCGCTGACCTCCCTGCTCTCTCTGCTCATCATATG 517  
RESULT 12  
AL450311 165110 bp DNA linear PRI 12-JUL-2001  
LOCUS Human DNA sequence from clone Rpl1-343J3 on chromosome 10, complete  
DEFINITION sequence.  
ACCESSION AL450311  
VERSION AL450311.11 GI:14626972  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 165110)  
AUTHORS Howden, P.  
TITLE Direct Submision  
JOURNAL Submitted (12-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
COMMENT On Jul 8, 2001 this sequence version replaced gi:14575291.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Sw:  
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 10, constructed by the Sanger Centre Chromosome 10  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr10  
Rpl1-343J3 is from the library RPl1-11.2 constructed by the group  
of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBACe3.6  
This sequence is the entire insert of clone Rpl1-343J3 The true  
left end of clone Rpl1-242G20 is at 139955 in this sequence. The  
true right end of clone Rpl1-404C6 is at 6588 in this sequence.  
FEATURES  
source  
1..165110  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/clone="Rpl1-343J3"  
/clone\_11b="RPl1-11.2"  
7..147  
/note="TH1LC repeat: matches 2..142 of consensus"  
151..293  
/note="MIR repeat: matches 2..148 of consensus"  
1136..1187  
/note="26 copies 2 mer gt 98% conserved"  
2960..3272  
/note="AluSg repeat: matches 1..310 of consensus"  
4144..4270  
/note="AluX repeat: matches 1..114 of consensus"  
4319..4494  
/note="AluX repeat: matches 118..293 of consensus"  
5662..6438  
/note="AluX repeat: matches 272..1095 of consensus"  
6916..8080  
/note="L1MEC repeat: matches 1168..2367 of consensus"  
9934..10087  
/note="MIR repeat: matches 25..184 of consensus"  
10353..10481  
/note="MIR repeat: matches 65..194 of consensus"  
11077..11372  
/note="AluX repeat: matches 1..294 of consensus"  
11600..11639  
/note="20 copies 2 mer tc 95% conserved"  
12479..12608  
/note="65 copies 2 mer at 90% conserved"  
13433..13876  
/note="MIRLC repeat: matches 1..466 of consensus"  
14372..14576

```
repeat_region /note="Alusq repeat: matches 1. .301 of consensus"
14577. .14633
/note="Alusq repeat: matches 243. .239 of consensus"
repeat_region 14843. .14943
/note="MER86 repeat: matches 9. .109 of consensus"
repeat_region 15404. .15613
/note="MIR repeat: matches 3. .213 of consensus"
repeat_region 16626. .16705
/note="MIR repeat: matches 60. .140 of consensus"
repeat_region 16928. .16993
/note="33 copies 2 mer gg 66% conserved"
repeat_region 17476. .17569
/note="MER81 repeat: matches 2. .114 of consensus"
repeat_region 17719. .18069
/note="LIMC4 repeat: matches 7617. .7977 of consensus"
repeat_region 18117. .18345
/note="MIR repeat: matches 7. .262 of consensus"
repeat_region 19169. .19400
/note="MIR repeat: matches 5. .239 of consensus"
repeat_region 19436. .19627
/note="LIMC5 repeat: matches 7728. .7917 of consensus"
repeat_region 19642. .19716
/note="MIR repeat: matches 48. .131 of consensus"
repeat_region 19969. .20010
/note="21 copies 2 mer tg 100% conserved"
repeat_region 20623. .20704
/note="41 copies 2 mer gt 85% conserved"
repeat_region 20738. .20773
/note="U2 repeat: matches 1. .36 of consensus"
repeat_region 20802. .20863
/note="MLT1J repeat: matches 1. .62 of consensus"
repeat_region 20889. .21263
/note="MLT1F repeat: matches 188. .541 of consensus"
repeat_region 21463. .21618
/note="MIR repeat: matches 46. .192 of consensus"
repeat_region 22019. .22326
/note="Alub repeat: matches 1. .306 of consensus"
repeat_region 22381. .22564
/note="U2 repeat: matches 2453. .2629 of consensus"
repeat_region 22836. .23174
/note="MLT1J repeat: matches 117. .413 of consensus"
repeat_region 23215. .23346
/note="MIR repeat: matches 48. .188 of consensus"
repeat_region 23388. .23532
/note="U2 repeat: matches 2097. .2230 of consensus"
repeat_region 23533. .23637
/note="Alub repeat: matches 1. .303 of consensus"
repeat_region 23838. .24137
/note="U2 repeat: matches 1754. .2097 of consensus"
repeat_region 24291. .24581
/note="Alusx repeat: matches 1. .300 of consensus"
repeat_region 24653. .24850
/note="MIR repeat: matches 1. .200 of consensus"
repeat_region 26189. .26685
/note="Cpg island"
/evidence=not_experimental
misc_feature 29453. .30918
/note="Cpg island"
/evidence=not_experimental
misc_feature 31084. .31156
/note="MIR repeat: matches 65. .138 of consensus"
repeat_region 31618. .31876
/note="Alu repeat: matches 29. .275 of consensus"
repeat_region 32767. .32830
/note="MIR repeat: matches 76. .139 of consensus"
repeat_region 33050. .33178
/note="43 copies 3 mer tcc 72% conserved"
misc_feature 35112. .36201
/note="Cpg island"
/evidence=not_experimental
repeat_region 36551. .36604
```

```
repeat_region /note="LIP15 repeat: matches 5480. .6157 of consensus"
38265. .38561
/note="Alusx repeat: matches 1. .300 of consensus"
repeat_region 39214. .39423
/note="MIR repeat: matches 3. .219 of consensus"
repeat_region 39489. .39601
/note="Alusx repeat: matches 1. .292 of consensus"
repeat_region 40050. .40189
/note="MIR repeat: matches 1. .144 of consensus"
repeat_region 41057. .41290
/note="MIR repeat: matches 8. .240 of consensus"
repeat_region 41296. .41460
/note="U2 repeat: matches 2569. .2730 of consensus"
repeat_region 41506. .41944
/note="U2 repeat: matches 1916. .2416 of consensus"
repeat_region 42388. .42698
/note="U2 repeat: matches 1448. .1779 of consensus"
repeat_region 44193. .44579
/note="TRIC repeat: matches 1. .371 of consensus"
repeat_region 44600. .44956
/note="LTR16A repeat: matches 90. .445 of consensus"
repeat_region 45240. .45300
/note="MER58A repeat: matches 37. .97 of consensus"
repeat_region 45798. .45909
/note="LTR41 repeat: matches 90. .192 of consensus"
repeat_region 46826. .46871
/note="23 copies 2 mer gt 97% conserved"
repeat_region 50684. .50969
/note="Alusq repeat: matches 1. .287 of consensus"
repeat_region 50980. .51291
/note="Alusx repeat: matches 1. .312 of consensus"
repeat_region 52222. .52519
/note="Alusx repeat: matches 1. .300 of consensus"
repeat_region 54065. .54260
/note="LIM4 repeat: matches 3865. .4055 of consensus"
repeat_region 54261. .54432
/note="PAM repeat: matches 2. .167 of consensus"
repeat_region 54433. .54629
/note="LIM4 repeat: matches 3652. .3865 of consensus"
repeat_region 54648. .54862
/note="LIR41 repeat: matches 11. .217 of consensus"
repeat_region 54863. .55236
/note="MLT1A1 repeat: matches 1. .365 of consensus"
repeat_region 55237. .55700
```

## Alignment Scores:

```
Pred. No.: 5.2e-28 Length: 165110
Score: 472.00 Matches: 90
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.68% Indels: 0
DB: 9 Gaps: 0
```

US-10-001-885-125 (1-121) x AL450311 (1-165110)

```
QY 32 GlnSerPheAsnArgIyHnIslgYAlaProProGlyProGlyProArgIngn 51
Db 90669 CAGAGCTTCTTCAACAGGGGCGATGCTCCCGCAGGGGGTCTGGCCCGCCAGAG 90128
QY 52 GlnAlaGlyAlaArgLeuGlyAlaAlaGlnSerProPheAsnArgLeuArgIngn 71
Db 90129 CAGGACAGGCGCAGGGTGGTCTGCTGCTCCCTTCAAGACTTCAACCGGAGCTG 90188
QY 72 ValAsnMetGlyPheProGlnTrpHisLeuGlyAsnHisAlaValGlnProValThrSer 91
Db 90189 GTCACATAGGCTTTCGACAGTGGCATCTTGGCAACATGCTGTGAGCGGTGACTCC 90248
QY 92 IleLeuIleuPheIleuIleuMetLeuLeuGlyValAlaGlyLeuLeuLeuValGlyLeu 111
Db 90249 ATCTGCTCTCTCTTCTGCTGCTCATGATGCTGTGGTGGCTCTCTGCTGCTGCTT 90308
QY 112 ValTyrLeuValSerHisLeuSerGlnArg 121
```

RESULT 13  
AC021954/c  
LOCUS  
DEFINITION  
AC021954  
AC021954  
AC021954.3 GI:7417809  
HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
1 (bases 1 to 173341)  
Britten, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 10, clone RP11-57E12  
Unpublished  
2 (bases 1 to 173341)

Britten, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Beckert, R., Beda, F.,  
Boguslavsky, L., Bouckhalter, B., Brown, A., Burkett, G., Castle, A.,  
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, K.,  
DeRellano, K., Dewar, K., Domino, M., Doyle, M., Fenebor, J.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyn, S., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howard, T., Johnson, R., Jones, C., Kamm, L., Karatas, A., Klein, J.,  
Lander, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
MachDonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, K.,  
McNeelers, R., Meldrum, J., Menus, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Olivar, T. M., Peterson, K.,  
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Testave, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.

REFERENCE  
JOURNAL  
AUTHORS  
Direct Submission  
Submitted (22-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 173341)

Britten, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavsky, L., Bouckhalter, B., Brown, A., Burkett, G.,  
Campoliano, A., Castle, A., Choepe, Y., Colangelo, M., Collins, S.,  
Collymore, A., Cooke, K., DeRellano, K., Dewar, K., Diaz, J. S.,  
Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyn, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howard, T., Johnson, R., Jones, C., Kamm, L., Karatas, A.,  
Klein, J., Lehoczy, K., Lamazares, R., Lander, T., Lehoczy, J.,  
Levine, R., Liu, G., Locke, K., Loh, G., Locke, K., MachDonald, P., Margulis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McNeelers, R.,  
Meldrum, J., Menus, L., Morrow, J., Miranda, C., Munge, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Olivar, T. M., Oliver, J., Peterson, K., Pierre, N.,  
Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Testave, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 5, 2000 this sequence version replaced gi:6984451.  
All repeats were identified using RepeatMasker:  
Smith, A. F. A. & Green, P. (1996-1997) RepeatMasker.html  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
Center: Whitehead Institute/MIT Center for Genome Research  
Center code: WIRB  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

## Project Information

Center project name: L5931  
Center clone name: 57 E.12

## Summary Statistics

Sequencing vector: M13, M77815, 100% of reads  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 161190 bases at least Q40  
Consensus quality: 166837 bases at least Q30  
Consensus quality: 168995 bases at least Q20  
Insert size: 176000; agarose-fp  
Insert size: 171041; sum-of-contigs  
Quality coverage: 3.7 in Q20 bases; sum-of-contigs  
Quality coverage: 3.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently  
consists of 24 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

|        |        |                               |
|--------|--------|-------------------------------|
| 1      | 1574   | contig of 1574 bp in length   |
| 1575   | 1674   | gap of 100 bp                 |
| 1675   | 3043   | contig of 1389 bp in length   |
| 3044   | 3143   | gap of 100 bp                 |
| 3144   | 5889   | contig of 2746 bp in length   |
| 5890   | 5989   | gap of 100 bp                 |
| 5990   | 8979   | contig of 2990 bp in length   |
| 8980   | 9079   | gap of 100 bp                 |
| 9080   | 13674  | contig of 4595 bp in length   |
| 13675  | 13774  | gap of 100 bp                 |
| 13775  | 18831  | contig of 5057 bp in length   |
| 18832  | 18931  | gap of 100 bp                 |
| 18932  | 23526  | contig of 4595 bp in length   |
| 23527  | 23626  | gap of 100 bp                 |
| 23627  | 27386  | contig of 3760 bp in length   |
| 27387  | 27486  | gap of 100 bp                 |
| 27487  | 32572  | contig of 5086 bp in length   |
| 32573  | 32672  | gap of 100 bp                 |
| 32673  | 38632  | contig of 5960 bp in length   |
| 38633  | 38732  | gap of 100 bp                 |
| 38733  | 43735  | contig of 5003 bp in length   |
| 43736  | 43835  | gap of 100 bp                 |
| 43836  | 49020  | contig of 5185 bp in length   |
| 49021  | 49120  | gap of 100 bp                 |
| 49121  | 53660  | contig of 4540 bp in length   |
| 53661  | 53760  | gap of 100 bp                 |
| 53761  | 59544  | contig of 5784 bp in length   |
| 59545  | 59644  | gap of 100 bp                 |
| 59645  | 66395  | contig of 6751 bp in length   |
| 66396  | 66495  | gap of 100 bp                 |
| 66496  | 74596  | contig of 8201 bp in length   |
| 74597  | 74796  | gap of 100 bp                 |
| 74797  | 81326  | contig of 6530 bp in length   |
| 81327  | 81426  | gap of 100 bp                 |
| 81427  | 89446  | contig of 8020 bp in length   |
| 89447  | 89546  | gap of 100 bp                 |
| 89547  | 96459  | contig of 6913 bp in length   |
| 96460  | 96559  | gap of 100 bp                 |
| 96560  | 106498 | contig of 9939 bp in length   |
| 106499 | 106598 | gap of 100 bp                 |
| 106599 | 116986 | contig of 10388 bp in length  |
| 116987 | 117086 | gap of 100 bp                 |
| 117087 | 128890 | contig of 11804 bp in length  |
| 128891 | 128990 | gap of 100 bp                 |
| 128991 | 147290 | contig of 18300 bp in length  |
| 147291 | 147390 | gap of 100 bp                 |
| 147391 | 173341 | contig of 25951 bp in length. |

## FEATURES

source  
1. 173341  
/organism="Homo sapiens"

/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10"  
/clone="RP11-57E12"  
/clone\_id="RP11-11 Human Male BAC"  
1.1574  
/note="assembly\_fragment"  
1675..3043  
/note="assembly\_fragment"  
3144..5889  
/note="assembly\_fragment"  
5990..8979  
/note="assembly\_fragment"  
9080..13674  
/note="assembly\_fragment"  
13775..18831  
/note="assembly\_fragment"  
18932..23526  
/note="assembly\_fragment"  
23627..27386  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left"  
27487..32572  
/note="assembly\_fragment"  
32673..38632  
/note="assembly\_fragment"  
38733..43735  
/note="assembly\_fragment"  
43836..49020  
/note="assembly\_fragment"  
49121..53660  
/note="assembly\_fragment"  
53761..59544  
/note="assembly\_fragment"  
59645..66395  
/note="assembly\_fragment"  
66486..74696  
/note="assembly\_fragment"  
74797..81326  
/note="assembly\_fragment"  
81427..89446  
/note="assembly\_fragment"  
89547..96459  
/note="assembly\_fragment"  
96560..106498  
/note="assembly\_fragment"  
106599..116986  
/note="assembly\_fragment"  
117087..128890  
/note="assembly\_fragment"  
clone\_end:P7  
vector\_side:right"  
128991..147290  
/note="assembly\_fragment"  
147391..173341  
/note="assembly\_fragment"

## Alignment Scores:

Pred. No.: 5,47e-28 Length: 173341  
Score: 472.00 Matches: 90  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 74.68% Indels: 0  
DB: 2 Gaps: 0

US-10-001-885-125 (1-121) x AC021954 (1-173341)

32 GlnSerPheAsnArgGlyHisGlyAlaProFrogGlyGlyProGlyProArgGlnGln 51  
CAGAGCTTCTTCATATAGGGGCGCATGTCCTCCCGCAGGGGGTCTGGCGCCCGCCACAG 161030

QY 52 GlnAlaGlyAlaAlaGlyGlyAlaAlaGlnSerProPheAsnArgGlyGln 71  
DB 161029 CAGGAGGTCGACGAGCTGGGCTGCTCAGTCCCTTCATGACCTCAACCGGAGCTG 160970  
QY 72 ValAsnMetGlyPheProGlnTrpHisLeuGlyAsnHisAlaValGluProValThrSer 91  
DB 160969 GTGAACATGGGCTTTCCCGCAGATGGCATCTCGGCACACATGCTGTGGAGCCGCTACCTCC 160910  
QY 92 IleLeuLeuLeuPheLeuLeuMetLeuGlyValArgGlyLeuLeuValGlyLeu 111  
DB 160909 ATCCGCTCCTCTCTCTCTCATGATGCTGTGCTGCTGCGCTCTCTCTGTTGGCCTT 160850  
QY 112 ValTyrLeuValSerHisLeuSerGlnArg 121  
DB 160849 GTCTACCTGGTGTCCCACTGATGTCAGCGG 160820  
RESULT 14  
AC127417 215050 bp DNA linear HTG 19-OCT-2002  
LOCUS AC127417/c  
DEFINITION Mus musculus chromosome UNK clone RP23-459W2, WORKING DRAFT  
SEQUENCE, 7 unordered pieces.  
AC127417  
AC127417.2 GI:24137619  
VERSION AC127417.2  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS McPherson,J.D. and Waterston,R.H.  
TITLE The sequence of Mus musculus clone  
JOURNAL Unpublished  
REFERENCE  
AUTHORS McPherson,J.D. and Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (19-OCT-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
On Oct 19, 2002 this sequence version replaced gi:21759524.  
COMMENT  
----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.edu  
Project information  
Center project name: M\_BA045M02  
----- Summary Statistics -----  
Sequencing vector: M13, 0%  
Sequencing vector: plasmid: 100%  
Chemistry: Dye-primer ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 212279 bases at least Q40  
Consensus quality: 213137 bases at least Q30  
Consensus quality: 213590 bases at least Q20  
Insert size: 19200; agarose-fp  
Insert size: 217944; sum-of-contigs  
Quality coverage: 11.89 in Q20 bases; agarose-fp  
Quality coverage: 10.54 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.



and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GZXS  
Center clone name: CH230-259G16  
----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 224747 bases at least Q40  
Consensus quality: 227981 bases at least Q30  
Consensus quality: 229752 bases at least Q20  
Estimated insert size: 228243; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 255979: contig of 255979 bp in length  
\* 255980 256079: gap of unknown length  
\* 256080 257349: contig of 1270 bp in length  
\* 257350 257449: gap of unknown length  
\* 257450 258815: contig of 1366 bp in length.

FEATURES  
source 1..258815  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-259G16"

misc\_feature

1..1609  
/note="wgs end extension  
clone\_end:5p6"

misc\_feature

/note="clone boundary  
complement(2865..3659)  
clone\_end:5p6"

misc\_feature

end\_sequence:B2169379"  
224098..225157  
/note="clone boundary  
clone\_end:17  
site:  
end\_sequence:B2169378"

ORIGIN

Alignment Scores:  
Pred. No.: 8,056-26 Length: 258815  
Score: 448.00 Matches: 85  
Percent Similarity: 95.56% Conservative: 1  
Best Local Similarity: 94.44% Mismatches: 4  
Query Match: 70.89% Indels: 0  
D3: 2 Gaps: 0

US-10-001-885-125 (1-121) x AC127817 (1-258815)

QY 32 GlnSerPhePheAsnArgGlyHisGlyValProProGlyGlyProGlyProArgGlnGln 51  
Db 220380 CAGGGCTTTTTCACAGAGGCCACAGTCACCTCCAGGGGGCCCTGGAGCCACCAGCAG 220321  
QY 52 GlnAlaGlyAlaArgLeuGlyValAlaGlnSerProPheAsnAspLeuAsnArgGlnLeu 71  
Db 220320 CAGCAGAGGCCGAGCTGGGCTGCCCAATCTCTTCAGTGAAGCTGAACCGGAGCTG 220261  
QY 72 ValAsnMetGlyPheProGlnTrpHisLeuGlyAsnHisAlaValGluProValThrSer 91  
Db 220260 GTGAACATGGGCTTCCACAAATGGCATTTGGGAACACGCTAGAGCCTGTGACCTCC 220201  
QY 92 IleLeuLeuLeuPheLeuLeuMetLeuGlyValAlaGlyLeuLeuValGlyLeu 111  
Db 220200 ATCCTCTGCTCTCTCTCTGCTCATGAGCTAGGGGTTCTGTGGCTCTCTGTGGCCTG 220141  
QY 112 ValTyrLeuValSerHisLeuSerGlnArg 121  
Db 220140 GTCTACCTGGTGTCTCACTGAGCCAGCGG 220111

Search completed: August 22, 2004, 09:14:41  
Job time : 2208 secs

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2004, 07:09:25 ; Search time 339 Seconds

(without alignments)  
1516.318 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632

Sequence: 1 MVRLLANGELVQDDPRVRR.....GVRGLLVGLVIVSHLSQR 121

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp  
-O=/gen2.1/USPTO.spool.p/US10001885/runat.17082004.151717.25892/app.query.fasta\_1.263  
-DB=N\_Geneseq 291n04 -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOFEFT=0 -UNITS=bits -START=1 -END=1 -MATRIX=dloume2 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10001885 @CGN 1 1 708 @runat.17082004.151717.25892 -NCPU=6 -ICPU=3  
-NO MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAUDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq 29jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001as:\*  
5: Geneseq2001bs:\*  
6: Geneseq2002s:\*  
7: Geneseq2003as:\*  
8: Geneseq2003bs:\*  
9: Geneseq2003cs:\*  
10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 632   | 100.0       | 365    | 7  | ABX92251 Human ova |
| 2          | 632   | 100.0       | 562    | 6  | ABZ11212 Human pol |
| 3          | 632   | 100.0       | 1049   | 2  | AAV59558 Human sec |
| 4          | 632   | 100.0       | 1049   | 6  | AB573545 Human cdn |
| 5          | 632   | 100.0       | 1049   | 8  | ACD82688 CDNA sequ |
| 6          | 632   | 100.0       | 1108   | 6  | ABL90057 Human pol |
| 7          | 632   | 100.0       | 1192   | 6  | ABN87820 Human ova |
| 8          | 577   | 91.3        | 550    | 4  | AAF82505 Human bre |

|   |    |       |      |      |   |                     |
|---|----|-------|------|------|---|---------------------|
| C | 9  | 434   | 68.7 | 1165 | 5 | AA64820 DNA encod   |
| C | 10 | 113   | 17.9 | 517  | 7 | ABX74673 Human CDN  |
| C | 11 | 110.5 | 17.5 | 1110 | 9 | ABE07153 Novel cod  |
| C | 12 | 110.5 | 17.5 | 2999 | 7 | AB557565 Human SEC  |
| C | 13 | 107   | 16.9 | 2668 | 7 | ACA03095 Human CDN  |
| C | 14 | 89    | 14.1 | 4177 | 5 | AA545077 CDNA down  |
| C | 15 | 89    | 14.1 | 4839 | 5 | ABV30270 Human encc |
| C | 16 | 88.5  | 14.0 | 2172 | 4 | ABV31456 Human pro  |
| C | 17 | 88.5  | 14.0 | 2172 | 7 | ACA19536 Pseudom    |
| C | 18 | 88.5  | 14.0 | 2172 | 8 | ACC59400 Microbial  |
| C | 19 | 88    | 13.9 | 1122 | 7 | ADA71003 Rice gene  |
| C | 20 | 86.5  | 13.7 | 494  | 8 | ACH43729 Human foe  |
| C | 21 | 86.5  | 13.7 | 1251 | 3 | AA62022 Hydrophob   |
| C | 22 | 86.5  | 13.7 | 2617 | 3 | AA62032 Hydrophob   |
| C | 23 | 86.5  | 13.7 | 2417 | 6 | ABX34832 Human CDN  |
| C | 24 | 86.5  | 13.7 | 2487 | 4 | AAH14983 Human CDN  |
| C | 25 | 86.5  | 13.7 | 2852 | 4 | AAH00603 Human mem  |
| C | 26 | 85    | 13.4 | 8031 | 4 | AAI07167 Human rep  |
| C | 27 | 85    | 13.4 | 1024 | 7 | ABZ83399 Toxicolog  |
| C | 28 | 84.5  | 13.4 | 1066 | 4 | AAI07167 Human rep  |
| C | 29 | 84.5  | 13.4 | 1170 | 7 | ACD05761 CDNA encc  |
| C | 30 | 84.5  | 13.4 | 1359 | 4 | AA94501 Human hyd   |
| C | 31 | 84.5  | 13.4 | 1484 | 6 | ABT11180 Human sec  |
| C | 32 | 84.5  | 13.4 | 1548 | 5 | AA582715 DNA encod  |
| C | 33 | 84.5  | 13.4 | 1550 | 2 | AA791346 Human HO-  |
| C | 34 | 84.5  | 13.4 | 1550 | 2 | AA791346 Human HO-  |
| C | 35 | 84.5  | 13.4 | 1550 | 2 | AA791346 Human hae  |
| C | 36 | 84.5  | 13.4 | 1550 | 3 | AA65120 Human hae   |
| C | 37 | 84.5  | 13.4 | 1550 | 3 | AA293230 Human Hem  |
| C | 38 | 84.5  | 13.4 | 1550 | 6 | ABE2886 Human hae   |
| C | 39 | 84.5  | 13.4 | 1550 | 6 | ABE2886 Breast ca   |
| C | 40 | 84.5  | 13.4 | 1550 | 6 | ABE2886 Breast ca   |
| C | 41 | 84.5  | 13.4 | 1550 | 6 | ABQ91974 Human NF-  |
| C | 42 | 84.5  | 13.4 | 1558 | 6 | ABN97191 Gene #368  |
| C | 43 | 84.5  | 13.4 | 1558 | 4 | AA94511 Human hyd   |
| C | 44 | 84.5  | 13.4 | 2259 | 7 | ADD29600 Human tun  |
| C | 45 | 84.5  | 13.4 | 2772 | 9 | ADA53012 Human cod  |
| C | 46 | 84.5  | 13.4 | 2772 | 9 | ADB63122 Human CDN  |

ALIGNMENTS

RESULT 1

ID ABX92251 standard; CDNA; 365 BP.

|    |                                                                       |               |  |  |  |
|----|-----------------------------------------------------------------------|---------------|--|--|--|
| AC | XX                                                                    | ABX92251;     |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| DT | 08-MAY-2003                                                           | (first entry) |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| DE | Human ovarian specific nucleic acid DEX0310_65.                       |               |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| KW | Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer; |               |  |  |  |
| KW | non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;     |               |  |  |  |
| KW | gynaecological.                                                       |               |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| OS | Homo sapiens.                                                         |               |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| PN | WO200292785-A2.                                                       |               |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| PD | 21-NOV-2002.                                                          |               |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| PF | 13-FEB-2002; 2002WO-US022271.                                         |               |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| PR | 13-FEB-2001; 2001US-0268290P.                                         |               |  |  |  |
| PR | 15-FEB-2001; 2001US-026834P.                                          |               |  |  |  |
| PA | (DIAD-) DIADEXUS INC.                                                 |               |  |  |  |
| PI | Salceda S, Macina R, Hu P, Recipon H, Karra K, Cafferty R;            |               |  |  |  |
| PI | Sun Y, Liu C,                                                         |               |  |  |  |
| XX | XX                                                                    |               |  |  |  |
| DR | WPI; 2003-120677/11.                                                  |               |  |  |  |
| DR | P-Psdb; ABU61063.                                                     |               |  |  |  |



XX New isolated OSNA nucleic acid and encoded polypeptide, useful for  
PT identifying, diagnosing, monitoring, staging, imaging and treating  
PT ovarian cancer and non-cancerous diseases in ovarian tissues.

PS Claim 1; Page 185-186; 224pp; English.

XX The invention relates to a new isolated nucleic acid, termed ovarian  
CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that  
CC encodes any of 53 fully defined protein sequences appearing as ABU61018-  
CC ABU61070 (termed ovarian specific proteins, OSP); (b) any of 76 fully  
CC defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a  
CC sequence having at least 60% sequence identity to the nucleic acid  
CC molecule of (a) or (b). Also included are a method for determining the  
CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector  
CC comprising an OSNA, a host cell comprising the vector, an isolated OSP  
CC polypeptide, an anti-OSP antibody or fragment, a method for determining  
CC the presence of an ovary specific protein in a sample and a vaccine  
CC comprising an OSP or OSNA. The methods and compositions of the present  
CC invention are useful for identifying, diagnosing, monitoring, staging,  
CC imaging and treating ovarian cancer and non-cancerous disease in ovary  
CC tissue. The present sequence is an OSNA of the invention

XX Sequence 365 BP; 60 A; 117 C; 108 G; 80 T; 0 U; 0 Other;

XX Alignment Scores:

|                        |          |               |     |
|------------------------|----------|---------------|-----|
| Pred. No.:             | 2,24e-51 | Length:       | 365 |
| Score:                 | 632.00   | Matches:      | 121 |
| Percent Similarity:    | 100.00%  | Conservative: | 0   |
| Best Local Similarity: | 100.00%  | Mismatches:   | 0   |
| Query Match:           | 100.00%  | Indels:       | 0   |
| DB:                    | 7        | Gaps:         | 0   |

US-10-001-885-125 (1-121) x ABX92251 (1-365)

QY 1 MetValArg11LeuAlaAsnGlyGluLeuValGlnAspAspProArgValArgThr 20  
DB 1 ATGGTGGCGATCTTGGCCAAATGGGGAATCGTGCAGATGACGCCCGAGTAGAAC 60  
QY 21 ThrThGlnProProArgGlySer11LeProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 61 ACTTACCAGCCACCAAGAGGTRGCTTCTTCACAGAGCTTCTTCAATAGGGGCCATGCT 120  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyValArgLeuGlyValAla 60  
DB 121 GGTCCCGCCGGGGGCTCTGCGCCCGCCGACAGACAGGAGGTGCGAGCTGGGTGCTGCT 180  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 181 CAGTCCCTCTTAAGTACCTCAACCGGAGCTGTGTGTAACATGGGCTTCCGCACTGGCAT 240  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSer11LeuLeuLeuPheLeuLeuMetMet 100  
DB 241 CTCGGCAACCATGCTGTGGAGCCGGTGAACCTCCATCTGCTCTTCTCTCATATG 300  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTrpLeuValSerHisLeuSerGln 120  
DB 301 CTTGGTGTTCGGGCTCTCTGTTGTGCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTT 360  
QY 121 Arg 121  
DB 361 CGG 363

RESULT 2  
AB211212  
ID AB211212 standard; cDNA; 562 BP.

AC AB211212;

DT 20-JAN-2003 (first entry)

KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
KW Parkinson's disease; Alzheimer's disease; autoimmune disease; diabetes;  
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KW arthritis; cytostatic; immunomodulator; nocitropic; neuroprotective;  
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
KW antiarthritic; gene; ss.

XX Homo sapiens.

XX WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US005095.

XX 05-MAR-2001; 2001US-00799451.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QH, Ren F,  
PI Xue HJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,  
PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2002-759812/82.

XX P-PSDB; ABP68995.

XX New polynucleotides comprising sequences assembled from expressed

PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
PT or coagulation disorders.

XX Claim 1; SEQ ID NO 94; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences (AB21119-  
CC AB21206) or their mature protein coding portion, active domain coding  
CC protein or complementary sequences. The polynucleotides are useful for  
CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell-  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, liver  
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIFO at ftp.wifo.int/pub/published\_pat\_sequences

XX Sequence 562 BP; 95 A; 176 C; 174 G; 117 T; 0 U; 0 Other;

XX Alignment Scores:

|                        |          |               |     |
|------------------------|----------|---------------|-----|
| Pred. No.:             | 3.73e-51 | Length:       | 562 |
| Score:                 | 632.00   | Matches:      | 121 |
| Percent Similarity:    | 100.00%  | Conservative: | 0   |
| Best Local Similarity: | 100.00%  | Mismatches:   | 0   |
| Query Match:           | 100.00%  | Indels:       | 0   |
| DB:                    | 6        | Gaps:         | 0   |

US-10-001-885-125 (1-121) x AB211212 (1-562)

QY 1 MetValArg11LeuAlaAsnGlyGluLeuValGlnAspAspProArgValArgThr 20  
DB 146 ATGGTGGCGATCTTGGCCAAATGGGGAATCGTGCAGATGACGCCCGAGTAGAAC 205  
QY 21 ThrThGlnProProArgGlySer11LeProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 206 ACTTACCAGCCACCAAGAGGTRGCTTCTTCACAGAGCTTCTTCAATAGGGGCCATGCT 265



Db 266 GCTCCCCAGGCGGCTCTCGGCCCCCGCCAGCAGCAGCAGGCTGCAGGCTGGCTGCTGCT 325  
Qy 61 GlnserProPheAnaapLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
Db 326 CAGTCCCCCTTCATGACCTTACACCGGACGCTGTGACATGGACCTTCGCGAGTGGCAT 385  
Qy 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuPheLeuMetMet 100  
Db 386 CTCGCAACCATGCTGTGGAGCCGATGACCTCCATCTGCTCTCCCTTCGCTCATCATG 445  
Qy 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValIleuValIleuSerHisLeuSerGln 120  
Db 446 CTTCGTGTCTGGCTCTCTCTCTGCTGTGGCTGTGACCTGTACCTGGTGTCCACTGATCAG 505  
Qy 121 Arg 121  
Db 506 CGG 508  
RESULT 3  
AAVS5558  
ID AAVS5558 standard; DNA, 1049 BP.  
AAVS5558;  
AC AAVS5558;  
XX  
XX  
DT 06-JAN-1999 (first entry)  
XX  
DE Human secreted protein gene 48 clone HECAT174.  
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; fetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
OS Homo sapiens;  
XX  
XX W09839448-A2.  
PD 11-SEP-1998.  
XX  
XX  
PF 06-MAR-1998; 98MO-US004493.  
XX  
PR 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040161P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043669P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0047492P.  
PR 23-MAY-1997; 97US-0047500P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.

PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047581P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.  
PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 13-JUN-1997; 97US-0048974P.  
PR 08-JUL-1997; 97US-0051926P.  
PR 16-JUL-1997; 97US-0052874P.  
PR 18-AUG-1997; 97US-0055724P.  
PR 22-AUG-1997; 97US-005630P.  
PR 22-AUG-1997; 97US-005631P.  
PR 22-AUG-1997; 97US-005632P.  
PR 22-AUG-1997; 97US-005635P.  
PR 22-AUG-1997; 97US-005637P.  
PR 22-AUG-1997; 97US-0056662P.  
PR 22-AUG-1997; 97US-0056664P.  
PR 22-AUG-1997; 97US-0056845P.  
PR 22-AUG-1997; 97US-0056862P.  
PR 22-AUG-1997; 97US-0056864P.  
PR 22-AUG-1997; 97US-0056872P.  
PR 22-AUG-1997; 97US-0056874P.  
PR 22-AUG-1997; 97US-0056875P.  
PR 22-AUG-1997; 97US-0056876P.  
PR 22-AUG-1997; 97US-0056877P.  
PR 22-AUG-1997; 97US-0056878P.  
PR 22-AUG-1997; 97US-0056879P.  
PR 22-AUG-1997; 97US-0056880P.  
PR 22-AUG-1997; 97US-0056881P.  
PR 22-AUG-1997; 97US-0056882P.  
PR 22-AUG-1997; 97US-0056884P.  
PR 22-AUG-1997; 97US-0056886P.  
PR 22-AUG-1997; 97US-0056887P.  
PR 22-AUG-1997; 97US-0056903P.  
PR 22-AUG-1997; 97US-0056908P.  
PR 22-AUG-1997; 97US-0056909P.  
PR 22-AUG-1997; 97US-0056910P.  
PR 22-AUG-1997; 97US-0056911P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057669P.  
PR 12-SEP-1997; 97US-0057761P.  
PR 02-OCT-1997; 97US-0058785P.  
PR 02-OCT-1997; 97US-0061060P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC;  
PI Bednarik DP, Endress GA, Yu G, Ni U, Feng P, Young PE, Greene JM;  
PI Ferrie AM, Duan R, Hu U, Florence KA, Olsen HS, Edner R, Brewer DA;  
PI Moore PA, Shi Y, Lafleur DM, Li Y, Zeng Z, Kyaw H;

XX WPI: 1998-506364/43.  
DR P-PSDB: AAW74778.

XX New isolated human genes and the secreted polypeptide(s) they encode -  
PT useful for diagnosis and treatment of e.g. cancers, neurological  
PT disorders, immune diseases, inflammation or blood disorders.

XX Claim 1, Page 280, 721pp; English.

XX This sequence represents a nucleic acid molecule designated Gene 48 from  
CC the human cDNA clone HFCAT174 (deposited as clone ATCC 97899 and ATCC  
CC 205045) which encodes a secreted human protein. The gene can be used to  
CC generate fusion proteins by linking to the gene to a human immunoglobulin  
CC Fc portion (e.g. AAV59502) for increasing the stability of the fused  
CC protein as compared to the human protein only. The invention relates to  
CC 186 novel genes and their fragments (nucleic acid sequences: AAV59511-  
CC V59812; amino acid sequences AAW74731-W75026) which are useful for  
CC preventing, treating or ameliorating medical conditions e.g. by protein  
CC or gene therapy. Also, pathological conditions can be diagnosed by  
CC determining the amount of the new polypeptides in a sample or by  
CC determining the presence of mutations in the new polynucleotides.  
CC Specific uses are described for each of the 186 polynucleotides, based on  
CC which tissues they are most highly expressed in (see AAV59511 for  
CC described uses).

XX Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 7.82e-51 Length: 1049  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-001-885-125 (1-121) x AAV59558 (1-1049)

QY 1 MetValAlaGlyIleuAlaAsnGlyLulIleValGlnAspAspProArgValArgThr 20  
DB 116 ATGGGCGGATCTGGCCAAATGGAATCTGAGCGAGCGACCCCGAGTGAAGACC 175  
QY 21 ThrThrglnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40  
DB 176 ACTACCCAGCCCAAGAGGTNGATTCTCGACAGAGCTTCTCAATAGGGGCGCATGT 235  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnIleValAlaGlyValArgGlyValAla 60  
DB 236 GCTCCCGCCAGGGGGTCTGGCCCCCGCAGACAGGAGGATGCGAGGCTGGGCTGCT 295  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnIlePhe 80  
DB 296 CAGTCCCTTCAAGACCTCAACCGGAGCTGGTGAACATGGGCTTCCGCAATGGCAT 355  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
DB 356 CTGGGCAACCAATGCTGTGAGCGGCTGACCTCCATCCCTGCTCTGCTCATATG 415  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
DB 416 CTGGGTTGTCGTGGCTCTCTGCTGTTGGCTTGTCTACCTGCTGCTGCTGCTGCTG 475  
QY 121 Arg 121

RESULT 4  
ID ABS73545  
AB573545 standard; cDNA; 1049 BP.

XX ABS73545;

XX 15-JAN-2003 (first entry)

XX Human cDNA #1 for novel secreted protein gene 48.

XX Human; ss; gene; secreted protein; autoimmune disease; chemotaxis;  
KW rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;  
KW liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;  
KW cerebrovascular disorder; cerebral ischemia; angiogenesis; sunburn;  
KW nervous system disorders; Alzheimer's disease; infection;  
KW ocular disorder; corneal infection; wound healing; tissue regeneration;  
KW epithelial cell proliferation; organ transplantation; food additive;  
KW preservative; nutritional.

XX Homo sapiens.

XX US6420526-B1.

XX 16-JUL-2002.

XX 08-SEP-1998; 98US-00149476.

XX 07-MAR-1997; 97US-0038621P.

XX 07-MAR-1997; 97US-0040161P.

XX 07-MAR-1997; 97US-0040162P.

XX 07-MAR-1997; 97US-0040163P.

XX 07-MAR-1997; 97US-0040333P.

XX 07-MAR-1997; 97US-0040334P.

XX 07-MAR-1997; 97US-0040336P.

XX 07-MAR-1997; 97US-0040626P.

XX 07-MAR-1997; 97US-0043311P.

XX 11-APR-1997; 97US-0043312P.

XX 11-APR-1997; 97US-0043313P.

XX 11-APR-1997; 97US-0043314P.

XX 11-APR-1997; 97US-0043315P.

XX 11-APR-1997; 97US-0043568P.

XX 11-APR-1997; 97US-0043569P.

XX 11-APR-1997; 97US-0043576P.

XX 11-APR-1997; 97US-0043578P.

XX 11-APR-1997; 97US-0043580P.

XX 11-APR-1997; 97US-0043669P.

XX 11-APR-1997; 97US-0043670P.

XX 11-APR-1997; 97US-0043671P.

XX 11-APR-1997; 97US-0043672P.

XX 11-APR-1997; 97US-0043674P.

XX 23-MAY-1997; 97US-0047492P.

XX 23-MAY-1997; 97US-0047500P.

XX 23-MAY-1997; 97US-0047501P.

XX 23-MAY-1997; 97US-0047502P.

XX 23-MAY-1997; 97US-0047503P.

XX 23-MAY-1997; 97US-0047581P.

XX 23-MAY-1997; 97US-0047582P.

XX 23-MAY-1997; 97US-0047583P.

XX 23-MAY-1997; 97US-0047584P.

XX 23-MAY-1997; 97US-0047585P.

XX 23-MAY-1997; 97US-0047586P.

XX 23-MAY-1997; 97US-0047587P.

XX 23-MAY-1997; 97US-0047588P.

XX 23-MAY-1997; 97US-0047589P.

XX 23-MAY-1997; 97US-0047590P.

XX 23-MAY-1997; 97US-0047592P.

XX 23-MAY-1997; 97US-0047593P.

XX 23-MAY-1997; 97US-0047594P.

XX 23-MAY-1997; 97US-0047595P.

XX 23-MAY-1997; 97US-0047596P.

XX 23-MAY-1997; 97US-0047597P.

XX 23-MAY-1997; 97US-0047598P.

23-MAY-1997; 97US-0047600P.  
 PR 23-MAY-1997; 97US-0047601P.  
 PR 23-MAY-1997; 97US-0047612P.  
 PR 23-MAY-1997; 97US-0047613P.  
 PR 23-MAY-1997; 97US-0047614P.  
 PR 23-MAY-1997; 97US-0047615P.  
 PR 23-MAY-1997; 97US-0047616P.  
 PR 23-MAY-1997; 97US-0047617P.  
 PR 23-MAY-1997; 97US-0047618P.  
 PR 23-MAY-1997; 97US-0047619P.  
 PR 06-JUN-1997; 97US-0048964P.  
 PR 06-JUN-1997; 97US-0048965P.  
 PR 13-JUN-1997; 97US-0049610P.  
 PR 08-JUL-1997; 97US-0051928P.  
 PR 16-JUL-1997; 97US-0051929P.  
 PR 18-AUG-1997; 97US-0055724P.  
 PR 22-AUG-1997; 97US-0056630P.  
 PR 22-AUG-1997; 97US-0056631P.  
 PR 22-AUG-1997; 97US-0056632P.  
 PR 22-AUG-1997; 97US-0056633P.  
 PR 22-AUG-1997; 97US-0056634P.  
 PR 22-AUG-1997; 97US-0056635P.  
 PR 22-AUG-1997; 97US-0056636P.  
 PR 22-AUG-1997; 97US-0056637P.  
 PR 22-AUG-1997; 97US-0056638P.  
 PR 22-AUG-1997; 97US-0056639P.  
 PR 22-AUG-1997; 97US-0056640P.  
 PR 22-AUG-1997; 97US-0056641P.  
 PR 22-AUG-1997; 97US-0056642P.  
 PR 22-AUG-1997; 97US-0056643P.  
 PR 22-AUG-1997; 97US-0056644P.  
 PR 22-AUG-1997; 97US-0056645P.  
 PR 22-AUG-1997; 97US-0056646P.  
 PR 22-AUG-1997; 97US-0056647P.  
 PR 22-AUG-1997; 97US-0056648P.  
 PR 22-AUG-1997; 97US-0056649P.  
 PR 22-AUG-1997; 97US-0056650P.  
 PR 22-AUG-1997; 97US-0056651P.  
 PR 22-AUG-1997; 97US-0056652P.  
 PR 22-AUG-1997; 97US-0056653P.  
 PR 22-AUG-1997; 97US-0056654P.  
 PR 22-AUG-1997; 97US-0056655P.  
 PR 22-AUG-1997; 97US-0056656P.  
 PR 22-AUG-1997; 97US-0056657P.  
 PR 22-AUG-1997; 97US-0056658P.  
 PR 22-AUG-1997; 97US-0056659P.  
 PR 22-AUG-1997; 97US-0056660P.  
 PR 22-AUG-1997; 97US-0056661P.  
 PR 22-AUG-1997; 97US-0056662P.  
 PR 22-AUG-1997; 97US-0056663P.  
 PR 22-AUG-1997; 97US-0056664P.  
 PR 22-AUG-1997; 97US-0056665P.  
 PR 22-AUG-1997; 97US-0056666P.  
 PR 22-AUG-1997; 97US-0056667P.  
 PR 22-AUG-1997; 97US-0056668P.  
 PR 22-AUG-1997; 97US-0056669P.  
 PR 22-AUG-1997; 97US-0056670P.  
 PR 22-AUG-1997; 97US-0056671P.  
 PR 22-AUG-1997; 97US-0056672P.  
 PR 22-AUG-1997; 97US-0056673P.  
 PR 22-AUG-1997; 97US-0056674P.  
 PR 22-AUG-1997; 97US-0056675P.  
 PR 22-AUG-1997; 97US-0056676P.  
 PR 22-AUG-1997; 97US-0056677P.  
 PR 22-AUG-1997; 97US-0056678P.  
 PR 22-AUG-1997; 97US-0056679P.  
 PR 22-AUG-1997; 97US-0056680P.  
 PR 22-AUG-1997; 97US-0056681P.  
 PR 22-AUG-1997; 97US-0056682P.  
 PR 22-AUG-1997; 97US-0056683P.  
 PR 22-AUG-1997; 97US-0056684P.  
 PR 22-AUG-1997; 97US-0056685P.  
 PR 22-AUG-1997; 97US-0056686P.  
 PR 22-AUG-1997; 97US-0056687P.  
 PR 22-AUG-1997; 97US-0056688P.  
 PR 22-AUG-1997; 97US-0056689P.  
 PR 22-AUG-1997; 97US-0056690P.  
 PR 22-AUG-1997; 97US-0056691P.  
 PR 22-AUG-1997; 97US-0056692P.  
 PR 22-AUG-1997; 97US-0056693P.  
 PR 22-AUG-1997; 97US-0056694P.  
 PR 22-AUG-1997; 97US-0056695P.  
 PR 22-AUG-1997; 97US-0056696P.  
 PR 22-AUG-1997; 97US-0056697P.  
 PR 22-AUG-1997; 97US-0056698P.  
 PR 22-AUG-1997; 97US-0056699P.  
 PR 22-AUG-1997; 97US-0056700P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057651P.  
 PR 05-SEP-1997; 97US-0057652P.  
 PR 05-SEP-1997; 97US-0057653P.  
 PR 05-SEP-1997; 97US-0057654P.  
 PR 12-SEP-1997; 97US-0058785P.  
 PR 02-OCT-1997; 97US-0061060P.  
 PR 06-MAR-1998; 98WO-US004493.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Rosen CA, Fiecher CL, Soppet DP, Carter KC,  
 PI Bednarik DR, Endress GA, Yu G, Ni U, Feng P, Young PE, Greene JM,  
 PI Perrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ender R, Brewer LA,  
 PI Moore PA, Shi Y, Latleir DW, Li Y, Zeng Z, Kyaw H,  
 XX  
 XX WPI; 2002-634736/68.  
 DR P-PSDB; ABG95227.  
 XX  
 XX New isolated human secreted protein for diagnosing, preventing, treating  
 PT or ameliorating medical conditions and used as a food additive or  
 XX preservative.  
 XX  
 XX Example 1; SEQ ID NO 58; 129pp; English.  
 PS  
 XX The invention relates to an isolated protein that is one of 186 human

CC secreted proteins, given in the specification, encoded by one of 309 cDNA  
 CC sequences also given in the specification. The protein is used in a  
 CC pharmaceutical composition used to prevent, treat or ameliorate a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angioneuromatosis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to  
 CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
 CC components. The present sequence represents a cDNA derived from a gene  
 CC encoding one of the novel human secreted proteins of the invention. Note:  
 CC This sequence did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=642056B1  
 XX  
 SO Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 7,82e-51 Length: 1049  
 Score: 632.00 Matches: 121  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0  
 US-10-001-885-125 (1-121) x ABG73545 (1-1049)  
 QY 1 MetValArgIleLeuAlaAsnGlyIleValGlnAspAspProArgValArgThr 20  
 DB 116 ATGTGCGGATTTGGCCATGGGAATGTTGCGAGACGACCCCGAGTGGAGAC 175  
 QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyIleGly 40  
 DB 176 ACTACCCAGCCCAAGAGGTGATTCCTCGACAGAGCTTCTTCAATAGGGGCGCATGT 235  
 QY 41 AlaProProGlyIleProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAla 60  
 DB 236 GCTCCCGCAGGGGGTCTGGCCCGCCGCGCAGCAGCGAGGCGCAGGCTGGGTCTCT 295  
 QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
 DB 236 CAGTCCCGCTTCAATGACTCTAACCAGCGAGCTGTGAGACATGGCTTCCCGAGTGGCAT 355  
 QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuPheLeuLeuMet 100  
 DB 356 CTCGGCAACCATGTGCTGTGAGCCGGGTGACCTCCATCCCTGCTCTCTCATGATG 415  
 QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
 DB 416 CTGTGATGTTGCGGCTCTCTCTGTTGGCTTGTCTTACCTGTGATGATGATG 475  
 QY 121 Arg 121  
 DB 476 CGG 478  
 RESULT 5  
 ID ACD82688 standard; cDNA; 1049 BP.  
 ID ACD82688;  
 AC ACD82688;  
 XX  
 XX 22-SEP-2003 (first entry)  
 DT  
 XX  
 XX cDNA sequence #48 containing coding region of a human secreted protein.  
 DE  
 XX



PA (MOORE/) MOORE P A.  
 PA (SHY/) SHY Y.  
 PA (LAFLEUR D W.) LAFLEUR D W.  
 PA (LIVY/) LIVY Y.  
 PA (ZENG Z.) ZENG Z.  
 PA (KYAW H.) KYAW H.  
 PI Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP,  
 PI Endress CA, Yu G, Ni J, Feng P, Young PB, Greene JM, Ferrie AM,  
 PI Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R,  
 PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
 DR WPI; 2003-521800/49.  
 DR P-PSDB; ABO34421.  
 XX New genes and its encoded prostate cancer antigen proteins, useful for  
 PT preventing, treating, ameliorating or diagnosing e.g. prostate cancers,  
 PT thymic hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral  
 PT ischemia.  
 PS Claim 4; SEQ ID NO 58; 260bp; English.  
 XX The present invention relates to the isolation of novel human secreted  
 CC proteins and the polynucleotide sequences encoding them. The invention  
 CC also discloses vectors, host cells, antibodies, and recombinant methods  
 CC for producing human secreted proteins. The polypeptide and polynucleotide  
 CC sequences for the secreted proteins are useful for preventing, treating,  
 CC ameliorating or diagnosing medical conditions such as hyperproliferative,  
 CC disorders (e.g. leukaemia or breast cancers), wounds, reproductive  
 CC disorders (e.g. haemophilia or  
 CC thrombocytopaenia), autoimmune disorders (e.g. graft-versus-host disease or  
 CC multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g. asthma),  
 CC viral or bacterial or fungal infections (e.g. AIDS or sepsis), renal  
 CC disorders (e.g. kidney failure), cardiovascular disorders (e.g. angina  
 CC pectoris, cerebral ischaemia or congenital heart defects), respiratory  
 CC disorders, neurological disorders (e.g. Alzheimer's disease or  
 CC Parkinson's disease), and inflammations (e.g. Crohn's disease). The  
 CC polynucleotide or polypeptide may also be used as vaccine adjuvants.  
 CC ACD82641-ACD82950 encode human secreted proteins or their fragments.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO web site at [seqdata.uspto.gov/patseq/seqid.htm](http://seqdata.uspto.gov/patseq/seqid.htm)  
 CC  
 XX  
 SQ Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 7.82e-51 Length: 1049  
 Score: 632.00 Matches: 121  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0  
 US-10-001-885-125 (1-121) x ACD82688 (1-1049)  
 QY 1 MetValArg11LeuAla1aAng1y1u1leVal1GlnASPaspAPProArgVal1ArgThr 20  
 Db 116 ATGTCGGGATCTTGGCAATGCGAAATGTGCACAGACGACCCCGAGTGGAGAC 175  
 QY 21 ThrThrGlnProProArgGlySer11eProArgGlnSerPhePheAsnArg1yHisGly 40  
 Db 176 ACTACCCAGCCACCAAGAGTACATCTCTCGACAGAGCTTCTCAATAGGGGCCATGCT 235  
 QY 41 AlaProProGly1y1yProGlyProArgGlnGlnAla1aGly1Val1ArgPheGly1a1a1a 60  
 Db 236 GCTCCCCCGAGGGGCTCTGGCCCCCGACGACGAGGACAGTCCAGGCTGGCTGCTGCT 295  
 QY 61 GlnSerProPheAsnArg1y1yProGlyProArgGlnGlnAla1aGly1Val1ArgPheGly1a1a1a 80  
 Db 296 CAGTCCCCCTTCATACCTCAACCGGACGCTGGTGAACATGAGGCTTTTCGACAGTGGCAT 355  
 QY 81 LeuGlyAsnHis1a1aVal1GlnProVal1ThrSer11eLeuLeuPheLeuLeuMetMet 100

Db 356 CTCCGGACACATGCTGTGAGCGCGGACCTTCATCTGCTCTCTGCTCATGAG 415  
 QY 101 LeuGlyVal1Arg1y1yLeuLeuVal1Gly1yLeuVal1SerHisPheSerGln 120  
 Db 416 CTTCGGTTCGTGGCTCTCTCTGCTTGTGCTTGTCTTACTGCTGTCTCCACCTGAGTCAG 475  
 QY 121 Arg 121  
 Db 476 CGG 478  
 RESULT 6  
 ID ABL90057 standard; cDNA, 1108 BP.  
 AC ABL90057;  
 XX 24-MAY-2002 (first entry)  
 DE Human polynucleotide SEQ ID NO 619.  
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200190304-A2.  
 PN 29-NOV-2001.  
 XX 18-MAY-2001; 2001WO-US016450.  
 PF 19-MAY-2000; 2000US-020551SP.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 PA Birze CE, Rosen CA,  
 PI WPI; 2002-122018/16.  
 DR P-PSDB; ABB89648.  
 XX  
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
 PT prevention of neural, immune system, muscular, reproductive,  
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
 PT disorders.  
 PS Claim 4; SEQ ID NO 619; 2081bp + Sequence Listing; English.  
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins  
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)  
 CC  
 SQ Sequence 1108 BP; 224 A; 274 C; 324 G; 271 T; 0 U; 5 Other;  
 Alignment Scores:  
 Pred. No.: 8.35e-51 Length: 1108

Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-001-885-125 (1-121) x ABL90057 (1-1108)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 147 ATGTTGGGAGATCTTGGCCAAATGGGAAATGTCACAGATGACGACCCCGAGTGGAGACC 206  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40  
DB 207 ACTACCCAGCCACCAAGAGGTACATTCCTCGACAGAGCTTCTTCAAYAGGGCCATGCT 266  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 267 GCTCCGCCAGGGGGTCTCTGGCCCCCGCAGACAGAGGAGGAGGCTGGGCTGCTGCT 326  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 327 CAGTCCCTTTCATATGACCTCAACCGGCAGCTGTGAAATGGGCTTTCGGCAGTGGCAT 386  
QY 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuLeuPheLeuLeuMetMet 100  
DB 387 CTGGGCAACCATGCTGTGGAGCCGGTGACCTCCATCTGCTCTTCCGTCATGATG 446  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
DB 447 CTGGTGTGTGGCTCTCTCTCTCTGTTGGCTTGTCTACTGCTGTCTCCACCTGACTCAG 506  
QY 121 Arg 121  
DB 507 CGG 509

RESULT 7

ABN87820 standard; cDNA; 1192 BP.

ABN87820;

12-NOV-2002 (first entry)

Human ovary specific nucleic acid SEQ ID NO:31.

Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;  
ovary specific gene; OSG; ovarian cancer; immune response; metastasis;  
chromosome 10; gene; ss.

Homo sapiens.

MO200240535-A2.

23-MAY-2002.

20-NOV-2001; 2001WO-US045011.

20-NOV-2000; 2000US-0252061P.

27-NOV-2000; 2000US-0253257P.

(DIAD-) DIADEXUS INC.

Salceda S, Macina RA, Recipon H, Caferkey R, Sun Y, Liu C;

WPI; 2002-471617/50.

New ovary specific genes and proteins, useful as a vaccine for treating

patients with ovarian cancer, or for diagnosing and monitoring the

presence and metastases of ovarian cancer in a patient.

Claim 1; Page 173-174; 260pp; English.

CC AEN87790 to AEN87882 represent human ovary specific nucleic acid (OSNA)  
CC sequences, and ABB79297 to ABB79370 represent ovary specific protein  
CC (OSP) sequences from the present invention. OSNA and OSP sequences have  
CC cytosolic activity, and can be used in vaccine production and gene  
CC therapy. An antibody that specifically binds to an OSP can be used for  
CC treating a patient with ovarian cancer, particularly for inducing an  
CC immune response against the ovarian cancer cell expressing the OSNA or  
CC OSP. The OSNAs and OSPs can also be used for diagnosing and monitoring  
CC the presence and metastases of ovarian cancer in a patient

Sequence 1192 BP; 262 A; 298 C; 349 G; 283 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 9,1e-51 Length: 1192  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-001-885-125 (1-121) x AEN87820 (1-1192)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 161 ATGTTGGGAGATCTTGGCCAAATGGGAAATGTCGACAGACGACCCCGAGTGGAGACC 220  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40  
DB 221 ACTACCCAGCCACCAAGAGGTACATTCCTCGACAGAGCTTCTTCAATAGGGCCATGCT 280  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 281 GCTCCGCCAGGGGGTCTCTGGCCCCCGCAGACAGAGGAGGCTGGGCTGCTGCT 340  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 341 CAGTCCCTTTCATATGACCTCAACCGGCAGCTGTGAAATGGGCTTTCGGCAGTGGCAT 400  
QY 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuLeuPheLeuLeuMetMet 100  
DB 401 CTGGGCAACCATGCTGTGGAGCCGGTGACCTCCATCTGCTCTTCCGTCATGATG 460  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
DB 461 CTGGTGTGTGGCTCTCTCTCTCTGTTGGCTTGTCTACTGCTGTCTCCACCTGACTCAG 520  
QY 121 Arg 121  
DB 521 CGG 523

RESULT 8

AAF82505 standard; cDNA; 550 BP.

AAF82505;

18-JUN-2001 (first entry)

Human breast tumour protein cDNA 15964.

Human; breast cancer; tumour; cytosolic; gene therapy; ss.

Homo sapiens.

MO200127276-A2.

19-APR-2001.

11-OCT-2000; 2000WO-US028255.

12-OCT-1999; 99US-00417031.

(CORI-) CORIXA CORP.

XX Harlocker SL, Dillon DC, Xu J;  
XX WPI, 2001-273773/28.  
XX  
XX New polypeptides encoded by polynucleotide sequences over-expressed in  
XX breast tumor tissue are useful to detect, monitor and treat breast  
XX cancer.  
XX  
XX Claim 13, Page 50, 52pp, English.  
XX  
XX The present sequence encodes a breast tumour-associated protein. It was  
XX shown to have at least two-fold overexpression in breast tumour tissue.  
XX The invention relates to an isolated polypeptide comprising at least an  
XX immunogenic portion of a breast tumor-specific protein, or its variant  
XX that retains the ability to react with antigen-specific antisera. The  
XX breast tumour polynucleotides, polypeptides and antibodies are useful for  
XX inhibiting development of breast cancer. The polynucleotides may be used  
XX to design primers and probes for detecting and monitoring breast cancer  
XX  
XX Sequence 550 BP; 102 A; 163 C; 164 G; 118 T; 0 U; 3 Other;  
XX  
XX Alignment Scores:  
XX Pred. No.: 6,13e-46 Length: 550  
XX Score: 577.00 Matches: 117  
XX Percent Similarity: 97.52% Conservative: 1  
XX Best Local Similarity: 96.69% Mismatches: 2  
XX Query Match: 91.30% Indels: 2  
XX DB: 4 Gaps: 0  
XX  
XX US-10-001-885-125 (1-121) x AAs64820 (1-550)  
XX  
XX QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
XX DB 115 ATGGTGGGATCTTGGCCATGGGAAATGCTGAGACAGACCCCGAGTGAAGACC 174  
XX QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyAsGly 40  
XX DB 175 ACTACCCAGCCACCAAGAGTACATCTCTGACAGAGCTTCTTCATAGGGGCGCATGCT 234  
XX QY 41 AlaProProGlyGlyProGlyProArgGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
XX DB 235 GCTCCGCCAGGGGCTCTGGCCCGCCGACAGCAGGACAGGTGCGAGGTGGTGGTCT 234  
XX QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
XX DB 295 CAGTCCCGCTTCATGACCTCAACCGGACGCTGTGAACATGGGCTTCCGACGTGGCAT 354  
XX QY 81 LeuGlyAsnHisAla-ValGlnProValThrSerIleLeuLeuPheLeuLeuMetMe 100  
XX DB 355 CTCGGCAACCATGCTGTGGAGCGCGGTGACCTCATCTGCT-CTTCTCGCTCAGAT 413  
XX QY 100 tLeuGlyValArgGlyLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerG1 120  
XX DB 414 GCTTGGTGGTGGTGGCTTCTGTGGTGGCTTGTCTACTGTGTGCCACCTGATCA 473  
XX QY 120 n 120  
XX DB 474 A 474  
XX  
XX RESULT 9  
XX AAs64820  
XX ID AAs64820 standard; cDNA; 1165 BP.  
XX AC AAs64820;  
XX DT 13-FEB-2002 (first entry)  
XX DB DNA encoding novel human diagnostic protein #624.  
XX XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KM food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX

OS Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI, 2001-639362/73.  
XX  
XX P-PSDB; ABB00633.  
XX  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX Claim 1, SEQ ID NO 624, 103pp, English.  
XX  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping.  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. AAs64197-AAs94564 represent novel human diagnostic  
XX coding sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1165 BP; 247 A; 279 C; 342 G; 297 T; 0 U; 0 Other;  
XX  
XX Alignment Scores:  
XX Pred. No.: 5,79e-32 Length: 1165  
XX Score: 434.00 Matches: 116  
XX Percent Similarity: 86.67% Conservative: 1  
XX Best Local Similarity: 85.93% Mismatches: 3  
XX Query Match: 68.67% Indels: 15  
XX DB: 5 Gaps: 3  
XX  
XX US-10-001-885-125 (1-121) x AAs64820 (1-1165)  
XX  
XX QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
XX DB 352 ATGGTGGGATCTTGGCCATGGGAAATGCTGAGACAGACCCCGAGTGAAGACC 411  
XX QY 20 rThrThrGlnProProArgGlySerIleProArgGlnSer-PhePheAsnArgGlyHisG 40  
XX DB 412 CACTACCCAGCCACCAAGAGTACATCTCTGACAGAGCTTCTTCATCAGGGGCGCATG 471  
XX QY 40 lYAlaProProGlyGlyPro-GlyProArgGlnGlnAlaGly-----AlaArgLeu 57  
XX DB 472 GTGCTCCCGCCAGGGGCTCGTGGCCCGC---CAGCAAGCAGGACAGGTGCGCAAGGCTG 528  
XX QY 58 G1YAlaAlaGln-SerProPheAsnAsp-LeuAsnArgGlnLeuVal-AsnMetGlyPhe 76

DB 529 GGTGCTGCTAGTCCCTCCCTTAAGCATTAACCGGACAGCTGTGTAACATGGGCTTT 588  
 QY 77 Progin-TrypHisLeu-GlyAsnHisAlaVal---GluProVal-ThrSerIleLeu-Le 94  
 DB 589 CCGCAGTGTGGACCTCTGGCAACCATGCTGTGGAGCCGGTAAGATCATCTCTCCT 648  
 QY 94 ULeuPheLeuLeuMetLeuGlyValArgGlyLeuLeuValGlyLeuValIle 114  
 DB 649 CCTCTTCCTGCTCATGATGCTGTGGTGTGGCCCTCCTCTGTGGCCCTGTCTACT 708  
 QY 114 UValSerHisLeu-SerGlnArg 121  
 DB 709 GTGTCCCACTGAGTCAAGCGG 731

RESULT 10  
 ABX74673/C  
 ID ABX74673 standard; cDNA; 517 BP.

AC ABX74673;  
 XX  
 DT 21-MAR-2003. (first entry)

DE Human cDNA sequence #128 differentially expressed in CC-RCC types.

KM Human; microarray; solid surface; immobilised probe; CC-RCC;  
 KM differential expression profile; aggressive CC-RCC tumour type;  
 KM non-aggressive CC-RCC tumour type; clear cell renal carcinoma;  
 KM gene expression profiling; tumour tissue; gene; se.

OS Homo sapiens.

PN MO200279411-A2.

PD 10-OCT-2002.

PF 29-MAR-2002; 2002WO-US009576.

PR 29-MAR-2001; 2001US-0279411P.

PA (VAND-) VAN ANDEL INST.

PI Haab B, Rhodes D, Teh BT, Takashi M;

DR WPI; 2003-040679/03.

PT New microarray, comprising a matrix of cDNA probe from a set of probes  
 PT immobilized to a solid surface in predetermined order, useful in the  
 PT prognosis of patients with clear cell renal carcinoma.

PS Claim 1; SEQ ID NO 459; 179bp; English.

XX The present invention relates to a microarray comprising a matrix of at  
 CC least one cDNA probe from a set of probes immobilised to a solid surface  
 CC in a predetermined order, where a row of pixels corresponds to replicates  
 CC of one distinct probe from the set. The probes are complementary to  
 CC nucleic acid sequences that are expressed differentially in aggressive as  
 CC compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)  
 CC and that hybridise to the probes under high stringency conditions. The  
 CC microarray is useful for the prognosis of patients with CC-RCC, wherein  
 CC aggressive and non-aggressive CC-RCC tumour types are characterised by  
 CC differential expression profiles of genes that hybridise with one or more  
 CC probes immobilised on the microarray. The arrays are useful for gene  
 CC expression profiling of tumour and normal tissues. The present sequence  
 CC represents a human cDNA sequence differentially expressed in aggressive  
 CC vs. non-aggressive type CC-RCC phenotypes

XX Sequence 517 BP; 166 A; 154 C; 90 G; 107 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.0705 Length: 517  
 Score: 113.00 Matches: 29  
 Percent Similarity: 96.67% Conservative: 0

Query Match: 17.88% Indels: 1  
 DB: 7 Gaps: 0  
 US-10-001-885-125 (1-121) x ABX74673 (1-517)

RESULT 11

ADE07153  
 ID ADE07153 standard; DNA; 1110 BP.

AC ADE07153;

DT 29-JUN-2004 (first entry)

DE Novel coding sequence (useful for identifying genetic disorders) #219.

KM novel gene; novel protein; tissue marker; molecular weight marker;  
 KM chromosome marker; genetic disorder; gene; ds.

OS Unidentified.

PN MO2003054152-A2.

PD 03-JUL-2003.

PF 10-DEC-2002; 2002WO-US039555.

PR 10-DEC-2001; 2001US-0339739P.

PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.

PR 14-MAR-2002; 2002US-0365384P.

PR 12-APR-2002; 2002US-0372381P.

PR 12-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao Q, Wang J;  
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

DR WPI; 2003-569235/53.

DR P-PsDB; ADE08064.

PT New polynucleotides, useful for expressing recombinant proteins for  
 PT analysis, characterization or therapeutic use, or as markers for tissues  
 PT in which the corresponding protein is preferentially expressed.

PS Claim 1; SEQ ID NO 219; 1177bp; English.

XX The invention comprises the amino acid and coding sequences of novel  
 CC proteins. The DNA and protein sequences of the invention are useful as:  
 CC markers for tissues in which the corresponding protein is preferentially  
 CC expressed; as molecular weight markers on gels; as chromosome markers or  
 CC tags; to identify chromosomes or to map related gene positions; and to  
 CC compare with endogenous DNA sequences in patients to identify potential  
 CC genetic disorders. The present DNA sequence represents a gene of the  
 CC invention.

XX Sequence 1110 BP; 330 A; 243 C; 344 G; 193 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.302 Length: 1110  
 Score: 110.50 Matches: 30  
 Percent Similarity: 50.00% Conservative: 17









Tue Aug 24 09:48:27 2004

us-10-001-885-125.p2n.rng

Page 14

Db           :   |||||  
          3689 TAGCCCTTGGGTCCCTGGAGCCCTTCTC 3717

Search completed: August 22, 2004, 08:37:48  
Job time : 347 secs

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2004, 07:19:21 ; Search time 2464 Seconds  
(without alignments)  
1466.446 Million cell updates/sec

Title: US-10-001-885-125  
Perfect score: 632  
Sequence: 1 MVRILANGELIVQDDPRVRT.....GVKGLLVGLVYLVSLSQR 121

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Rgapop 6.0, Rgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seg length: 0  
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=x1p  
-Q=/cgn2\_1/USPRO.spool/p/US1001885/runat.17082004.151718.25914/app.query.fasta\_1.263  
-DB=EST -QMT=fastap -SUFFIX=p2n.txt -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=Blasum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US1001885.GCGN.1.1.5180 @runat.17082004.151718.25914 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_esthum:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pla:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_piro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 632   | 100.0       | 505    | 14    | CA942977 1r66f05.y |
| 2          | 632   | 100.0       | 566    | 14    | CB296571 12B2027   |
| 3          | 632   | 100.0       | 688    | 12    | BI521387 603081177 |
| 4          | 632   | 100.0       | 713    | 12    | BI193763 602948418 |
| 5          | 632   | 100.0       | 714    | 10    | BE568253 602184629 |
| 6          | 632   | 100.0       | 765    | 13    | BK095713 BK095713  |
| 7          | 632   | 100.0       | 980    | 10    | BE904502 601494571 |
| 8          | 632   | 100.0       | 1104   | 12    | BC331398 602433369 |
| 9          | 618   | 97.8        | 528    | 10    | AM960624 EST372695 |
| 10         | 602   | 95.3        | 1059   | 13    | BK420823 BX420823  |
| 11         | 592   | 93.7        | 447    | 9     | AA452397 4x29404.r |
| 12         | 574.5 | 90.9        | 454    | 10    | AM462809 BP230010B |
| 13         | 574.5 | 90.9        | 468    | 10    | BE237464 146771 MA |
| 14         | 574.5 | 90.9        | 523    | 10    | BE664644 152244 MA |
| 15         | 574.5 | 90.9        | 537    | 10    | BF046179 BP250022B |
| 16         | 574.5 | 90.9        | 569    | 10    | BE588875 194533 BA |
| 17         | 574.5 | 90.9        | 694    | 14    | CB437353 684811 MA |
| 18         | 570.5 | 90.3        | 527    | 9     | AV591976 AV591976  |
| 19         | 570.5 | 90.3        | 653    | 13    | BU671222 N16C-1706 |
| 20         | 568.5 | 90.0        | 600    | 14    | CA527226 8035-48 M |
| 21         | 568.5 | 90.0        | 651    | 10    | BE628297 BE628297  |
| 22         | 568.5 | 90.0        | 783    | 14    | CK129164 AGENCOURT |
| 23         | 568.5 | 90.0        | 784    | 12    | BI658578 603284311 |
| 24         | 568.5 | 90.0        | 823    | 12    | BG662915 602797603 |
| 25         | 568.5 | 90.0        | 832    | 10    | BR161467 601770455 |
| 26         | 568.5 | 90.0        | 891    | 12    | BI905000 603169151 |
| 27         | 568.5 | 90.0        | 949    | 13    | BU511854 AGENCOURT |
| 28         | 568.5 | 90.0        | 1095   | 11    | AK008338 Mus muscu |
| 29         | 568.5 | 90.0        | 1732   | 11    | AK036802 833705 MA |
| 30         | 567.5 | 89.8        | 546    | 14    | CP181245 B18054 MA |
| 31         | 566.5 | 89.6        | 658    | 14    | CP181245 B18054 MA |
| 32         | 564.5 | 89.3        | 553    | 13    | BU088757 SNEST4a89 |
| 33         | 564.5 | 89.3        | 558    | 10    | BE074572 222023 MA |
| 34         | 562.5 | 89.0        | 526    | 12    | BX523683 BX523683  |
| 35         | 542.5 | 85.8        | 519    | 29    | CG607245 OST285976 |
| 36         | 542   | 85.8        | 631    | 12    | BM714109 UI-E-ET0- |
| 37         | 540.5 | 85.5        | 691    | 13    | BY708262 BY708262  |
| 38         | 539.5 | 85.4        | 1069   | 12    | BA414577 602986159 |
| 39         | 537.5 | 85.0        | 443    | 9     | AA428101 2w49802.r |
| 40         | 537.5 | 85.0        | 504    | 29    | CG607398 OST286267 |
| 41         | 529.5 | 83.8        | 350    | 10    | BE237465 146772 MA |
| 42         | 524   | 82.9        | 1027   | 12    | BI522220 603081177 |
| 43         | 523.5 | 82.8        | 474    | 9     | AA032373 m144b06.r |
| 44         | 513.5 | 81.2        | 497    | 29    | CG648434 OST40197  |
| 45         | 507.5 | 80.3        | 505    | 9     | AA271091 va71b04.r |

ALIGNMENTS

RESULT 1  
CA942977  
LOCUS CA942977 505 bp mRNA linear EST 30-DEC-2002  
DEFINITION ttr6f05.y1 HR85 islet Homo sapiens CDNA clone IMAGE:607449 5',  
CA942977  
ACCESSION CA942977 GI:27431457  
VERSION CA942977.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 505)

## AUTHORS

Melton, D., Brown, J., Keny, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Birstein, A., Schmidt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V., Williams, T., Jackson, Y., and Bowers, Y.

## TITLE

## JOURNAL

Unpublished (2000)  
Endocrine Pancreas Consortium

## COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Harvard University, Howard Hughes Medical Institute  
Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel.: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@molp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@mol.wustl.edu)

Seg primer: -400P from Glibco  
High quality sequence stop: 463.

## FEATURES

## source

Location/Qualifiers

1..505  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6607449"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_lib="HR85 islet"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; CDNA made by oligo-dT priming;  
Size selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110. E-mail: hinoue@imgate.wustl.edu, Tel.:  
314-362-1916, Fax: 314-747-2692."

## ORIGIN

## Alignment Scores:

| Pred. No.:             | 1..946-45 | Length:       | 505 |
|------------------------|-----------|---------------|-----|
| Score:                 | 632.00    | Matches:      | 121 |
| Percent Similarity:    | 100.00%   | Conservative: | 0   |
| Best Local Similarity: | 100.00%   | Mismatches:   | 0   |
| Query Match:           | 100.00%   | Indels:       | 0   |
| DB:                    | 14        | Gaps:         | 0   |

US-10-001-885-125 (1-121) x CA942977 (1-505)

QY 1 MetValArgGileuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 123 ATGGTGGGATTTTGGCCATATGGGAAATCGTGCAGATGACGACCCCGCATGACACC 182  
QY 21 ThrThGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 183 ACTACCCAGCCACCAAGAGTGCATTCCTGCAGACGACTTCTTCAACAGGGCCATGCT 242  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 243 GCTCCCGCCAGGGGGCTTGGCCCCCGCCAGCAGCAGGAGGCCAGGGCTGGGCTCT 302  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTyrHis 80  
DB 303 CAGTCCCCCTTCAATGACTCAACCGGACGCTGGGAAATGAGGCTTCCGACATGGCAT 362  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuLeuPheLeuLeuMetMet 100  
DB 363 CTGGGCAACCAAGCTGTGAGCGCGGTGACCTTCATCTGCTCTTCTCTGCTCATGATG 422

## DB

423 CTGGATTTCTGGCTGCTCTCTGTTGGCTTGTCTACCTGGTGGTCCACCTGATGACG 482

## QY

121 Arg 121

## DB

483 CCG 485

## RESULT 2

## CB296571

## LOCUS

CB296571 566 bp mRNA linear EST 28-FEB-2003  
12822027.rev\_1\_B02\_r\_025.ab1 Chimpanzee brain library Xcos Pan  
troglodytes cDNA clone 12822027\_rev\_1\_B02\_r\_025.ab1 5', mRNA  
sequence.

## DEFINITION

## ACCESSION

CB296571  
CB296571.1 GI:28622001

## KEYWORDS

## SOURCE

EST.  
Pan troglodytes (chimpanzee)  
Pan troglodytes

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

Selection on human genes as revealed by comparisons to chimpanzee  
cDNA  
Genome Res. (2003) In press.  
Contact: Paabo S  
Evolutionary Genetics  
Max-Planck-Institute for evolutionary Anthropology  
Deutscher Platz 6, 04103 Leipzig, Germany  
Tel: +49-(0)-341-3550 500  
Fax: +49-(0)-341-3550 555  
Email: paabo@eva.mpg.de  
Seq primer: M13 reverse.

## JOURNAL

## COMMENT

## FEATURES

## source

1..566  
/organism="Pan troglodytes"  
/mol\_type="mRNA"  
/db\_xref="taxon:9598"  
/clone="12822027\_rev\_1\_B02\_r\_025.ab1"  
/sex="male"  
/tissue\_type="Brain, presumably cortex"  
/dev\_stage="adult"  
/lab\_host="Bpicanian Coli (TM) XL-10-Gold"  
/clone\_lib="Chimpanzee Brain library Xcos"  
/note="Vector: pUC19; Site\_1: SfiI-B; The  
library was prepared using the SMART cDNA library  
construction kit (Clontech), doing only primer extension,  
but not PCR amplification of the cDNA. The only deviation  
from the published protocol was that we cloned the cDNA  
into a plasmid vector."

## ORIGIN

## Alignment Scores:

| Pred. No.:             | 2..286-45 | Length:       | 566 |
|------------------------|-----------|---------------|-----|
| Score:                 | 632.00    | Matches:      | 121 |
| Percent Similarity:    | 100.00%   | Conservative: | 0   |
| Best Local Similarity: | 100.00%   | Mismatches:   | 0   |
| Query Match:           | 100.00%   | Indels:       | 0   |
| DB:                    | 14        | Gaps:         | 0   |

US-10-001-885-125 (1-121) x CB296571 (1-566)

QY 1 MetValArgGileuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 169 ATGGTGGGATTTTGGCCATATGGGAAATCGTGCAGATGACGACCCCGCATGACACC 228  
QY 21 ThrThGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 229 ACTACCCAGCCACCAAGAGTGCATTCCTGCAGACGACTTCTTCAACAGGGCCATGCT 288

Db 289 GCTCCCCAGGGGCTCTGCGCCCCCGACAGCAGCAGCTGCGAGCTGGGCTGCT 348

Qy 61 GlnserProPheAsnAAspleuAsnArgGlnLeuValAsnMetGlyPheProGlnTPHs 80

Db 349 CAGTCCCCCTTCATGACCTCAACCGGAGCTGTGAACATGGGCTTCCACAGTGGCAT 408

Qy 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuPheLeuMetMet 100

Db 409 CTGGGCAACCATGCTGTGAGCGCGGTGACCTCCATCTGCTCTCTCTGCTCATGATG 468

Qy 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120

Db 469 CTGGGTGTGTGTGGCTCTCTCTGCTGTGTGGCTGTGTCTTCACTGGTGTCCACCTGAGTCAG 528

Qy 121 Arg 121

Db 529 CGA 531

RESULT 3

LOCUS B1521387 688 bp mRNA linear EST 29-AUG-2001

DEFINITION 603081177F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:520570 5',

ACCESSION B1521387

VERSION B1521387.1 GI:15346179

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 688)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabds-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM1554 row: 1 column: 19  
High quality sequence stop: 688.

FEATURES

source

1. 688

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5220570"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_120"

/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 2.99e-45 Length: 688

Score: 632.00 Matches: 121

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-10-001-885-125 (1-121) x B1521387 (1-688)

Qy 1 MetValArgIleLeuAlaAsnGlyLulleValGlnAspAspProArgValArgThr 20

Db 140 ATGTGTGGGATCTTGGCCATGGGAAATCGTGACGATGACAGCCCGGATGGAGACC 199

Qy 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40

Db 200 ACTACCCAGCCACCAAGAGTGTGATCTCCACAGAGCTTCTTCAATAGGGGCCATGTGT 259

Qy 41 AlaProProGlyGlyProGlyProArgGlnGlnAlaGlyValArgLeuGlyValAla 60

Db 260 GCTTCCCCAGGGGCTCTGCGCCCCCGACAGCAGCAGGTCAGAGCTGGTGTCTGT 319

Qy 61 GlnserProPheAsnAAspleuAsnArgGlnLeuValAsnMetGlyPheProGlnTPHs 80

Db 320 CAGTCCCCCTTCATGACCTCAACCGGAGCTGTGAACATGGGCTTCCGACAGTGGCAT 379

Qy 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuPheLeuMetMet 100

Db 380 CTGGGCAACCATGCTGTGAGCGCGGTGACCTCCATCTGCTCTCTCTGCTCATGATG 439

Qy 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120

Db 440 CTGGGTGTGTGTGGCTCTCTCTGCTGTGTGGCTGTGTCTTCACTGGTGTCCACCTGAGTCAG 499

Qy 121 Arg 121

Db 500 CGG 502

RESULT 4

LOCUS B1193763 713 bp mRNA linear EST 10-JUL-2001

DEFINITION 602948418F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:5091412 5',

ACCESSION B1193763

VERSION B1193763.1 GI:14648783

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 713)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabds-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM1855 row: d column: 05  
High quality sequence stop: 709.

FEATURES

source

1. 713

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5091412"

/lab\_host="DH10B (phage-resistant)"

/tissue\_type="epithelial carcinoma cell line"

/clone\_lib="NIH\_MGC\_42"

/note="Organ: pancreas; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN Note: this is a NIH\_MGC Library. |"

## Alignment Scores:

Score: 3.14e-45 Length: 713  
 Percent Similarity: 632.00 Matches: 121  
 Best Local Similarity: 100.00% Conservative: 0  
 Query Match: 100.00% Mismatches: 0  
 DB: 12 Indels: 0  
 Gaps: 0

US-10-001-885-125 (1-121) x B1193763 (1-713)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
 DB 31 ATGTGGCGATCTTGGCCATGGGGAATCGTGCAGATGACACCCCGAGTGAGACC 90  
 QY 21 ThrThgInPProProArgGlySerIleProArgGlnSerPhePheAsnArgGly 40  
 DB 91 ACTACCCAGCCACCAAGAGTACATTCCTCGACAGACTTCTTCAATAGGGGCCATGGT 150  
 QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
 DB 151 GGTCCCGCAGGGGGGCTCTGGCCCCCGCAGCAGCAGGAGTGCAGGCTGGGCTGCT 210  
 QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnValGlyPheProGlnTrpHis 80  
 DB 211 CAGTCCCCCTTCAATGACCTCAACCGGCGAGCTGTGAAACATGGGCTTCCGCACTGGCAT 270  
 QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
 DB 271 CTCGGCAACCATGCTGTGAGCCGGTGACCTCCATCTCTCTCTCTCTCTCTCTCTCT 330  
 QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValIleValIleValSerHisLeuSerGln 120  
 DB 331 CTGGGTGTTGGGCTCTCTCTGTTGGCTGTCTACCTGCTGCTCCACCTGAGTCAG 390  
 QY 121 Arg 121  
 DB 391 CGG 393

RESULT 5  
 BFs68253 714 bp mRNA linear EST 12-DEC-2000  
 LOCUS 602184629F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:4300518 5',  
 DEFINITION mRNA sequence.  
 BFs68253  
 ACCESSION BFs68253.1 GI:11641633  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 714)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cga@bbs-remail.nih.gov

Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L10CM159 row: j column: 07  
 High quality sequence stop: 712.  
 Location/Qualifiers  
 1..714

FEATURES  
 source /organism="Homo sapiens"

/clone="IMAGE:4300518"  
 /tissue type="epithelioid carcinoma cell line"  
 /lab host="DH10B (phage-resistant)"  
 /clone lib="NIH MGC 42"  
 /note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT printing;  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCCAGG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library. |"

## Alignment Scores:

Score: 3.14e-45 Length: 714  
 Percent Similarity: 632.00 Matches: 121  
 Best Local Similarity: 100.00% Conservative: 0  
 Query Match: 100.00% Mismatches: 0  
 DB: 10 Indels: 0  
 Gaps: 0

US-10-001-885-125 (1-121) x BFs68253 (1-714)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
 DB 153 ATGTGGCGATCTTGGCCATGGGGAATCGTGCAGATGACACCCCGAGTGAGACC 212  
 QY 21 ThrThgInPProProArgGlySerIleProArgGlnSerPhePheAsnArgGly 40  
 DB 213 ACTACCCAGCCACCAAGAGTACATTCCTCGACAGACTTCTTCAATAGGGGCCATGGT 272  
 QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
 DB 273 GGTCCCGCAGGGGGGCTCTGGCCCCCGCAGCAGCAGGAGGAGTGCAGGCTGGGCTGCT 332  
 QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnValGlyPheProGlnTrpHis 80  
 DB 333 CAGTCCCCCTTCAATGACCTCAACCGGCGAGCTGTGAAACATGGGCTTCCGAGTGCAT 392  
 QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
 DB 393 CTCGGCAACCATGCTGTGAGCCGGTGACCTCCATCTCTCTCTCTCTCTCTCTCTCT 452  
 QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValIleValIleValSerHisLeuSerGln 120  
 DB 453 CTGGGTGTTGGGCTCTCTCTGTTGGCTGTCTACCTGCTGCTCCACCTGAGTCAG 512  
 QY 121 Arg 121  
 DB 513 CGG 515

RESULT 6  
 BX095713 765 bp mRNA linear EST 04-FEB-2003  
 LOCUS BX095713 Soares fetal N2HFe\_9w Homo sapiens cDNA clone  
 DEFINITION IMAGE998B081942; IMAGE:787879, mRNA sequence.  
 BX095713  
 ACCESSION BX095713  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 765)  
 Ebert L., Heil O., Hennig S., Neubert P., Partsch E., Peters M.,  
 Radelof U., Schneider D. and Korn B.  
 Human Unigenesec - R2PD3  
 TITLE Unpublished (2003)  
 JOURNAL Contact: ina.roffs

COMMENT



RZPD; IMAGE998B081942.  
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
Human UnigeneSet - RZPD (RZPDLIB No.972)  
http://www.rzpd.de/clonecards/cgi-  
bin/showlib.pl.cgi?response=11&no=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de  
This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
M13r, Primer sequence: TTTCACACGAGAAACAGCTATAC.  
Location/Qualifiers

## FEATURES

Source

1..765  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE998B081942 ; IMAGE.787879"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares, total\_fetus\_Nb2H8\_9w"  
/note="Vector: pRT3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from pooled 8-9 week  
(total) fetus material with a Not I - oligo(dt) primer [5'  
TGTTACCACTGACGAGGAGGCGGCTATTTTCTTTTCTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pRT3D vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,46e-45 Length: 765  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-001-885-125 (1-121) x BX095713 (1-765)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 45 ATGGTGGGAGATCTTGGCCAAATGGGGAATCTGTGAGATGACGACCCCGAGTGAAGACC 104  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 105 ACTACCCAGCCACCAAGAGGTAGCATTCCTGACAGAGCTTCTTCAACAGGAGCCATGGT 164  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAla 60  
DB 165 GCTCCCCAGGGGGTCTCTGGCCCCCGCCAGCAGCAGCAGGCTGGCAGGCTGGGTGCT 224  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 225 CAGTCCCCCTTCATGACCTCAACCGGACAGCTGTGAACATGGGCTTCCGAGTGGCAT 284  
QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
DB 285 CTGACCAACATGCTGTGAGAGCCGTGACCTCAACCTGCTCTTCTTCTGCTCATATG 344  
QY 101 LeuGlyValArgGlyLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
DB 345 CTGAGTGTGCTGGGCTCTCTCTGTTGGCCTTGTCACTGCTGTCCACCTGAGTCAG 404  
QY 121 Arg 121  
DB 405 CGG 407

RESULT 7

BE904302  
LOCUS BE904302 980 bp mRNA linear EST 20-OCT-2000  
DEFINITION 601494571P2 NIH\_MGC\_70 Homo sapiens cDNA clone IMAGE.3896656 5',  
RNA sequence.  
ACCESSION BE904302  
VERSION BE904302.1 GI:10396415  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 980)  
NIH-MGC http://mhc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: LHM9689 row: n column: 17  
High quality sequence stop: 711.  
Location/Qualifiers

## FEATURES

source

1..980  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE.3896656"  
/tissue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 70"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."

## ORIGIN

## Alignment Scores:

Pred. No.: 4,88e-45 Length: 980  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-001-885-125 (1-121) x BE904302 (1-980)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 140 ATGGTGGGAGATCTTGGCCAAATGGGGAATCTGTGAGATGACGACCCCGAGTGAAGACC 199  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 200 ACTACCCAGCCACCAAGAGGTAGCATTCCTGACAGAGCTTCTTCAATGGGAGCCATGGT 259  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAla 60  
DB 260 GCTCCCCAGGGGGTCTCTGGCCCCCGCCAGCAGCAGCAGGCTGGGCTGGGTGCT 319  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 320 CAGTCCCCCTTCATGACCTCAACCGGACAGCTGTGAACATGGGCTTCCGAGTGGCAT 379  
QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
DB 380 CTGACCAACATGCTGTGAGAGCCGTGACCTCAACCTGCTCTTCTTCTGCTCATATG 439  
QY 101 LeuGlyValArgGlyLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120

Db 440 CTGAGTGTTCGAGGCTCTCTCTGTTGGCTGTCTACTGTATGCCACCTGATGAC 499

QY 121 Arg 121

Db 500 CGG 502

RESULT 8  
BG331398

LOCUS BG331398

DEFINITION 1104 bp mRNA linear EST 27-FEB-2001  
602433369P1 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:4550896 5',  
mRNA sequence.

ACCESSION BG331398

VERSION BG331398.1 GI:13137937

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1104)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NIH-MGC <http://mgs.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: DCTD/BTP/Gazdar  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LNCM1244 row: 5 column: 17  
High quality sequence, stop: 636.  
Location/Qualifiers

FEATURES  
source 1..1104  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4550896"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH MGC 18"  
/note="Organ: lung; Vector: pOT87; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

ORIGIN

Alignment Scores:  
Pred. No.: 5,76e-45 Length: 1104  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-10-001-885-125 (1-121) x BG331398 (1-1104)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20

Db 152 ATGGTGGGATCTTGGCCATGCGAAATGTGTCAAGATGACACCCCGAGTGGAGACC 211

QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40

Db 212 ACTACCAAGACACCAAGAGGTAGCATTCCTCGACAGAGCTTCTCAATAGGGGCGCATGCT 271

QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgGlnGlyAlaAla 60

Db 272 GCTCCCGCAGGGGCTCTGGCCCCCGCCGACGACGAGGCGAGGCTGGGTGCTGCT 331

QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80

Db 332 CAGTCCCTCTTAATGACTCTCAACCGCAGCTGTGTGATGATGGGCTTCCGAGTGGCAT 391

QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuMetMet 100

Db 392 CTCGGAAACCATGCTGTGGAGCCGGTGAACCTCCATCCTGCTCTTCTGCTATGATG 451

QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120

Db 452 CTGAGTGTTCGAGGCTCTCTCTGTTGGCCCTGTCTACTGTGTCCACCTGATGACG 511

QY 121 Arg 121

Db 512 CGG 514

RESULT 9

LOCUS AW960624

DEFINITION 528 bp mRNA linear EST 01-JUN-2000  
EST372695 MAGE resequences, MAGP Homo sapiens cDNA, mRNA sequence.

ACCESSION AW960624

VERSION AW960624.1 GI:8150308

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 528)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Hedge, P., Qi, R., Abernathy, K., Dharwad, S., Gaspard, R., Gay, C.,  
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, F.J. and  
Quackenbush, J.  
Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
Unpublished (2000)  
Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: [johnq@igf.org](mailto:johnq@igf.org)  
Plate: 148  
Seq primer: Reverse.

FEATURES  
source 1..528  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_id="MAGE resequences, MAGP"  
/note="Vector: pBluescriptSKm"

ORIGIN

Alignment Scores:  
Pred. No.: 3,45e-44 Length: 528  
Score: 618.00 Matches: 118  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.78% Indels: 0  
DB: 10 Gaps: 0

US-10-001-885-125 (1-121) x AW960624 (1-528)

QY 4 IleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThrTrpThrGln 23

Db 3 ATCTTGGCAATGGGAAATGTGTCAAGATGACACCCCGAGTGGAGACCATACCCAG 62

QY 24 ProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGlyAlaProPro 43

Db 63 CCACCAAGAGGTAGCATTCCTCGACAGAGCTTCTCAACAGGGGCGCATGCTGCCCA 122

QY 44 GlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgGlnGlyAlaAlaGlnSerPro 63

Db 123 GGGGCTCTGGCCCCCGCCGACGACGAGGCGAGGCTGGGTGCTGCTGCTGCTGCT 182